Human Progress, Medicine and AI

Sep 11, 2025

The Parker Library, Corpus Christi College, Cambridge University

STIFEL | Healthcare

STIFEL | Healthcare

Table of Contents

Executive Summary

Executive Summary

We spend an enormous of time in the life sciences industry thinking about the sources of innovation and measures that can influence it such as government pricing policies and patents.

In this report we take innovation as a given and instead focus on a much less discussed topic: the **adoption** of innovation.

Sometimes great ideas don't make it to practice in a timely way.

The average time from bench to bedside of medical innovations today runs over thirty years.

The pace of innovation adoption, in our view, is just as important an area to explore as the drivers of innovation itself.

The big question we focus on in this report is why medical innovations are so slow to be adopted.

And how might this all change in the future?

To facilitate this exploration, we will start with a study of progress in a specific area of women's health and the esoteric field of naval medicine.

We show that tens of millions of preventable deaths occurred when solutions to medical problems were known but ignored, resisted, or poorly transmitted.

Count us as concerned and very interested in what our society can do in the future to make sure that past errors in innovation adoption are not repeated.

This is obviously incredibly relevant when current forces in Washington DC do not always favor the advancement of science for the benefit of patients.

We then argue that today's confluence of AI with the open science movement represents a profound change in how knowledge is disseminated and adopted in clinical practice.

The conversation quickly turns into a discussion of what is going on in the field of medicine and AI itself. You will see on the next several pages comments that came in from various readers of a draft of this report. Universally readers focus on what going on in AI today. Is it good? Is it bad? What should we be doing about AI? And why isn't AI itself not being used as much as perhaps it should?

We discuss these topics in some depth in this report.

Comments from Readers on This Report

Tess Cameron, Managing Director, RA Capital

Forceps were also hidden for over 100 years by the Chamberlens, including by blindfolding women who were in labor. Maybe if patents had been more established the Chamberlens would have more quickly disseminated this technology :) Incentives!

We simply cannot afford NOT to use AI to make healthcare more accessible/available. The MSFT dx cases are great but we need more real-world testing and training of docs, as you mention, to avoid things like automation bias (some studies have shown AI increases doc confidence in a dx even if it does not increase accuracy). The Topol editorial you highlight outlines some of the paths for adoption well.

As I'm sure you saw on your trip to China, the use of AI in hospitals for basic tasks (like summarizing patient health state and medical record before going into a meeting) seemed pretty impressive at least at the superficial level that I observed and is being actively encouraged.

From an Experienced Physician

I just had a lengthy discussion about AI and medicine with a young physician yesterday. He just finished his Residency and reviewed some of the ways he used AI as a resident.

Interestingly, at his institution, there is little support for using AI tools due to cost constraints for staff dermatologists. Hence, work-arounds are required.

He takes pictures of skin lesions and relies on AI for diagnostic suggestions. Sometimes there would be useful suggestions that he had failed to consider.

But this is all done with his personal phone in an "off the grid" way.

He also pointed out that in the derm space AI was more useful/accurate for "white skin lesions" than for lesions arising in skin with more melanin. Presumably, this reflects the datasets that were used to train AI.

Comments from Readers on This Report

Sami Yusuf, Chair Corp Development, Mayo Clinic

We at the Mayo Clinic are investing in AI for medicine in a huge way. Despite all of this, you won't be surprised to hear that physicians across the U.S. are often resisting bringing AI into their medical practices. Adoption of new technologies in medicine has been slower, in general, that one might like.

What is so interesting is that China has been far more receptive to using AI in medicine as a way to deliver better care to patients that don't have the great access to begin with.

China has introduced the first "AI Hospital" which treats patients with computers with only a single physician overseeing the activities.

No one in the U.S. is anywhere near to this type of idea.

While this is still very much experimental it's becoming increasingly clear that AI can make a major difference in getting innovation and good care to patients.

Tal Zaks, Partner, OrbiMed

The one tool you should absolutely mention is Open Evidence. It has replaced Uptodate as the medical resident's "go to", and the shift is telling. Open Evidence is an AI-LLM approach - just ask the question and it will give you the answer. Which immediately brings to mind the challenge of **trust** - it recommends something, and we all intuit that AI is better than an MD, but how do you know in the vast universe of use cases that it will be so for the specific question you are asking?

As it relates to broader adoption, I agree with the direction of where we're heading - especially AI's ability to discern patterns in data that lead to predictors so strong one would be wise to take action - but there is an infrastructure challenge everywhere you look. It's great that AI-pathology can determine a Gleason score with an accuracy that has the good Dr. Gleason turning in his grave; it is non-sensical that we use AI to determine a Gleason score - in itself an imperfect predictive marker rather than come up with an "AI score" directly predictive of relevant mortality) outcomes. But neither will matter much until we are able to digitize pathology slides on a more routine and consistent basis (few pathology slides ever get computerized today).

Comments from Readers on This Report

Jeremy Norman on AI Being So Good in Medicine

I would say that a primary value that AI will have in medicine may be wading through the immense information overload, and hopefully separating out misinformation from valid information, and then preparing some kind of summary for human consumption. A generalization that recently caught my attention is that AI operates in some ways that humans can understand, and also in ways that humans cannot understand. This is fundamentally different from computing until now. This paradox of humans consulting information systems that are partly beyond human comprehension, if correct, suggests that humans may eventually be consulting machines that retain a kind of oracular presence or Syblline mystery. The fear that computers become smarter than people is a very old concern, going back to the first electronic processing in 1945, but the reality of this concern was not something to worry about until now.

Note: Jeremy Norman is an expert on the history of science, medicine, an author of a website on humans and information www.historyofinformation.com, and an accompanying book on the history of computing (with Diana Hook). He is an author of a forthcoming book on the second printing revolution (via the Grolier Club) which reviews the mechanization of printing.

Jeremy Norman on Delays in Innovation Adoption

I was intrigued by your historical study of the delays of adoption of key medical discoveries throughout history.

I have not seen such a review before.

As you know, part of the adoption of medical innovation today comes from the medical establishment including physicians, part from government, and part from society as a whole.

It is reasonable to assume, I think, that AI may result in much faster cures to currently incurable diseases. Such cures will be well worth all the effort going into AI.

Whether the medical establishment and government will adopt those cures in a timely manner is less certain, though we saw it happen with the adoption of the mRNA vaccines.

For that to happen in the current environment in the US seems less likely. As you know well, other countries are presently more open to innovation.

Comments from Readers on This Report

Jeremy Norman on Information Manipulation/AI

I read your essay and did not notice the "politically incorrect" references to the particular problem of misinformation that we face today. I think part of the pervasiveness of what might be called social biases has to do with information overload, caused by the flood of free information, and society's reaction to it. It could also be that social biases are built-into humanity and can be manipulated using whatever communication tools are available. In the past the flood of scientific information was limited by education, and the cost of books, newspapers, subscriptions, to only a fraction of society. Today anyone with an internet connection faces the flood of scientific and pseudoscientific information, and the potential for this information to be manipulated. Some people face the overload and bias challenge head on and try to sort out scientific truth from falsehood, and other people--perhaps most of society--just adopt a bias--perhaps one promoted by a "leader"-- and may close their mind to the "actual" facts if they can find them. This issue was not something that I thought would become as pervasive as it has become. It suggests that some of the misinformation problems that we are currently facing may remain pervasive even after political changes.

Elena Viboch, Managing Director, General Catalyst

We need policy solutions that make research more efficient. Changes in funding incentives could be an incredible catalyst for leveraging AI to help researchers do more with less. It is high time for any researcher, anywhere, to be able to analyze their experimental data in the context of results from all prior publicly funded research. What would happen if the NIH and other federal funding organizations withheld future grants from any PI who had not complied with study reporting requirements? A wealth of data would immediately become available. What would happen if more detailed methods reporting requirements were instituted for future grants? More labs would be empowered to allocate their budget to advancing science, and would spend less time working to re-establish assays. What would happen if a replicability bonus were applied to future grant funding? If scientists could get their funding expanded when two independent labs successfully replicated their results, how much more progress would future funding yield? And how is it that researchers within a pharmaceutical company still are not able to take advantage of the wealth of experimental and clinical translational data owned by their employer? Not only should it be easily searchable and indexed, analyzing multi-modality data in the context of the literature should also be de-rigueur. AI can fix that.

Comments from Readers on This Report

Elliot Hershberg, Partner, Amplify Bio

I liked the historical view on (slow) medical adoption. I think the case for AI adoption given a constant rate of knowledge generation is clear.

But LLMs aren't just doing search and synthesis, in the next five to ten years they could start to really restructure how medical evidence is created. Consider systematic reviews.

Several leaders in the field recently published their efforts to automate systematic review generation with large language models.

You could imagine a point where the volume of creation and frequency of updates simply *requires* the use of AI models to be able to keep up with it.

Acknowledgements

In addition to the comments shared here I received comments on earlier version of this report from Charlotte Blease, Rob Carlson, Brandon Chang, Andrew Dannenberg, Masaki Doi, Alec Grant, Nick Haddad, Peter Kolchinsky, Susan Lewis, Ernest Li, Stephanie Leouzon, Kishan Patel and Travis Read.

Thank you all for the time to help improve this report.

Tim Opler
Managing Director
Healthcare Investment Banking Team
Stifel
(oplert@stifel.com)

Our Cover Page:
Matthew Parker Preserves His Manuscripts

In 1574, the year before his death, Archbishop of Canterbury Matthew Parker donated his library of about 480 manuscripts and about 1000 printed books to the library of Corpus Christi College, Cambridge.

Parker was an avid book collector, salvaging medieval manuscripts dispersed at the dissolution of the monasteries. He was particularly intent on preserving materials relating to Anglo-Saxon England, motivated by his search for evidence of an ancient English-speaking Church independent of Rome. Parker's extraordinary collection of documents remains at Corpus Christi College, and consists of items spanning from the sixth-century Gospels of St. Augustine to sixteenth century records relating to the English Reformation. (Source: Jeremy Norman, The History of Information)

Progress in Managing Maternal Mortality

Maternal Death in Childbirth Was a Huge Problem in Tudor England

England's King Henry VIII Lost Two Wives in Childbirth

Henry VIII (28 June 1491 – 28 January 1547) was King of England from 1509 until his death in 1547. Henry is known for his six marriages and his efforts to have his first marriage (to Catherine of Aragon) annulled. His disagreement with Pope Clement VII about such an annulment led Henry to initiate the English Reformation.

Henry VIII was a **miscreant** of his day. Henry brought radical changes to the Constitution of England, ushering in the theory of the divine right of kings in opposition to papal supremacy. He frequently used charges of treason to quell dissent, and those accused were often executed. Henry achieved many of his political aims through his chief ministers, some of whom were executed when they fell out of his favor. We know, this all sounds **vaguely familiar.**

In addition, Henry executed two of his wives, Anne Boleyn and Catherine Howard, both based on accusations of adultery.

But, just as interestingly (to us at least), Henry lost two wives in childbirth – both to **puerperal fever** (today called post-partum septicemia).

Jane Seymour (1508–1537) – Queen of England, third wife of Henry VIII (and his favorite). Died just days after giving birth to the future Edward VI.

Catherine Parr (1512–1548) – Queen of England, sixth wife of Henry VIII. Died from complications of childbirth. She outlived Henry by a year and six months.

The Terrible Toll of Puerperal Fever

While no one kept the statistics, well over 100 million women have died from post-partum septicemia (puerperal fever).* This is on the same order of magnitude as the count of humans who have died from war. Some of the more famous cases of death from puerperal fever are listed below. Even Henry VIII's mom died of the disease. Wikipedia keeps a list of famous women who died in childbirth – it runs on for many pages.

Elizabeth of York (1466–1503): The wife of King Henry VII and mother of Henry VIII, died of a post-partum infection a week after giving birth to her eighth child.

Lucrezia Borgia (1480–1519): Infamous Italian Noblewoman, illegitimate daughter of Pope Alexander VI and former governor of Spoleto died shortly after giving birth to her tenth child in 1519 at the age of 39.

Empress Isabella of Portugal (1503–1539) – Wife of Charles V, Holy Roman Emperor. Died of puerperal fever after the birth of her sixth child.

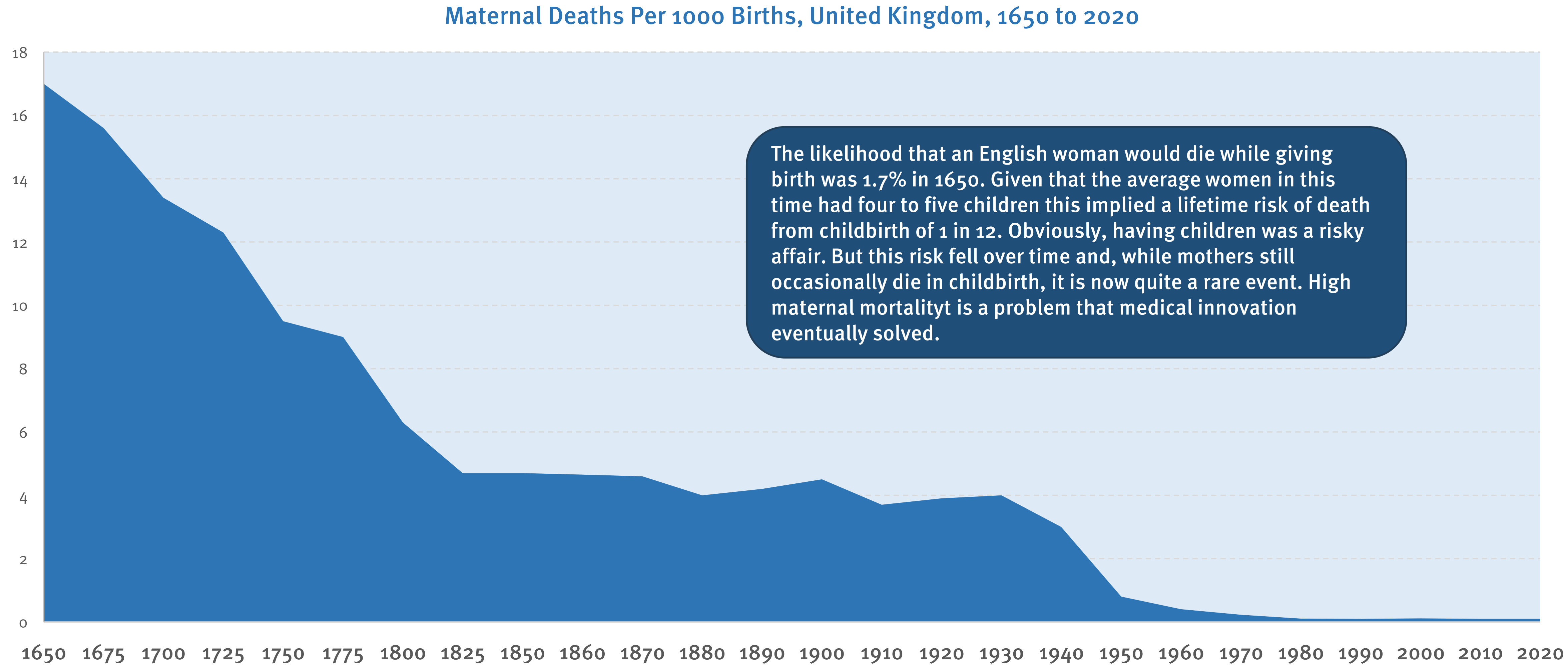
Émilie du Châtelet (1706–1749): A French mathematician and physicist known for her translation of and commentary on Isaac Newton's *Principia*. She died a few days after giving birth to her daughter.

Martha Jefferson (1748–1782): Thomas Jefferson's wife died at the age of 33 of puerperal fever, a fate shared by one of her daughters. She never lived to see her husband become President.

Mary Wollstonecraft (1759–1797): The English writer and early feminist philosopher died of sepsis ten days after giving birth to her second daughter, Mary Shelley, who would later write *Frankenstein*.

* Obtained by multiplying the number of births by the pre-1850 death rate from this condition: Cumulative births › 100 billion through 1850 times a death rate of 0.8% = roughly 400 million deaths. See https://www.prb.org/articles/how-many-people-have-ever-lived-on-earth/. Depending on the era up to half of these deaths have been linked to puerperal fever. Sources for death rates from puerperal fever are provided on the next two pages.

In 1650 an English Woman Had a One in Twelve Chance of Dying in Childbirth



Sources: (data for 1650 to 1800, Table 6.29 from Wrigley, Davies, Oeppen and Schofield, English Population History from Family Reconstitution, 1580–1837, Cambridge University Press, 1997); for 1800 to 1970 (Irvine Loudon, Death in Childbirth, Clarendon Press, Oxford, 1992), 1980 to 2020 (WHO Mortality Database).

Like England, Other Major Countries Saw a Big Drop in Maternal Mortality in the 1930s

Italy

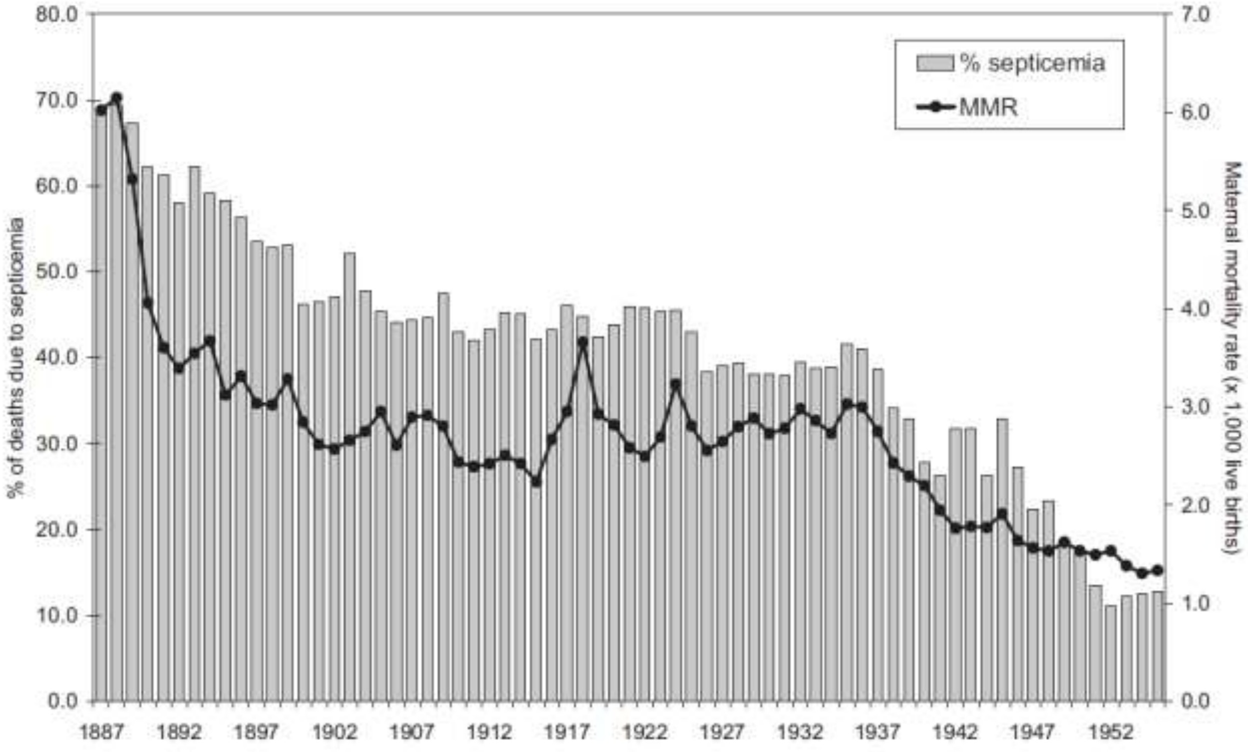
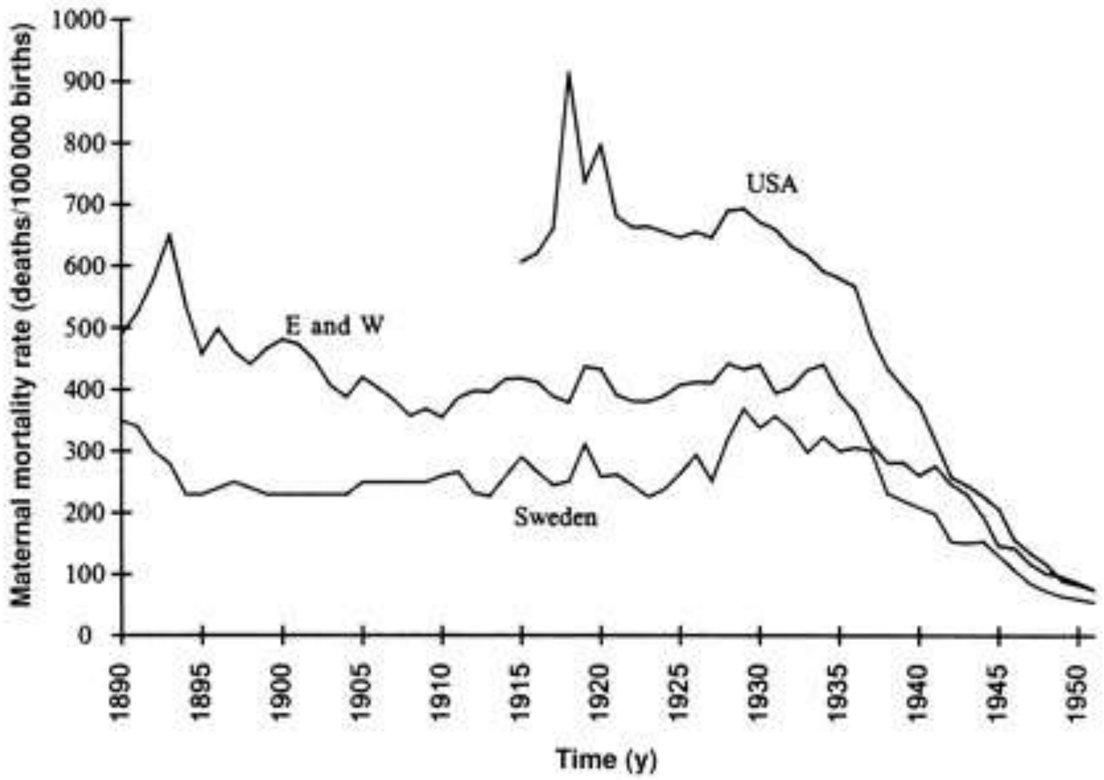


Fig. 1 Maternal mortality rate and per cent of maternal deaths for septicaemia in Italy, 1887–1955.

US, England and Wales, Sweden



Sources: Matteo Manfredini et al., Maternal Mortality in 19th- and Early 20th-century Italy, *Social History of Medicine*, 2019 and Irvine Loudon, "Maternal mortality in the past and its relevance to developing countries today," *The American Journal of Clinical Nutrition*, 2000, Volume 72, Issue 1, 241S - 246S.

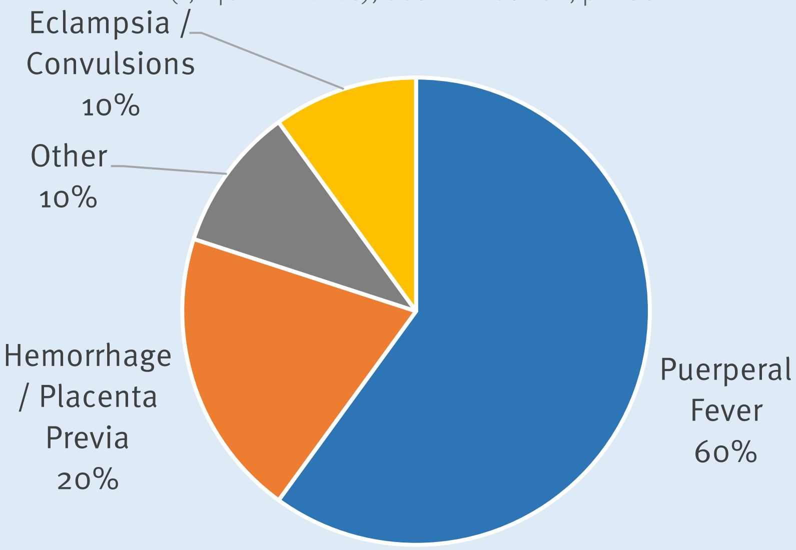
Causes of Maternal Mortality: Early 19th & 20th Centuries

Medicine tends not to keep great statistics about its failures. There are no organized data that we can find on the causes of maternal mortality from the 16th to 18th Centuries but anecdotal reports indicate that puerperal fever, obstructed labor and post-partum hemorrhage were the most important causes of maternal death in childbirth in this time. Oxford Professor Irvine Loudon did a nice job of finding later data on the causes of maternal mortality. What emerges from his data (shown below) is that the main causes of death in childbirth in the 1800s were puerperal fever followed by post-partum hemorrhage, obstructed labor and eclampsia. By the 1900's, death rates had dropped a lot and while puerperal fever was still with us, eclampsia had become more important in a relative sense.

US, Early 19th Century

Main Causes of US Maternal Mortality 1811 - 1838

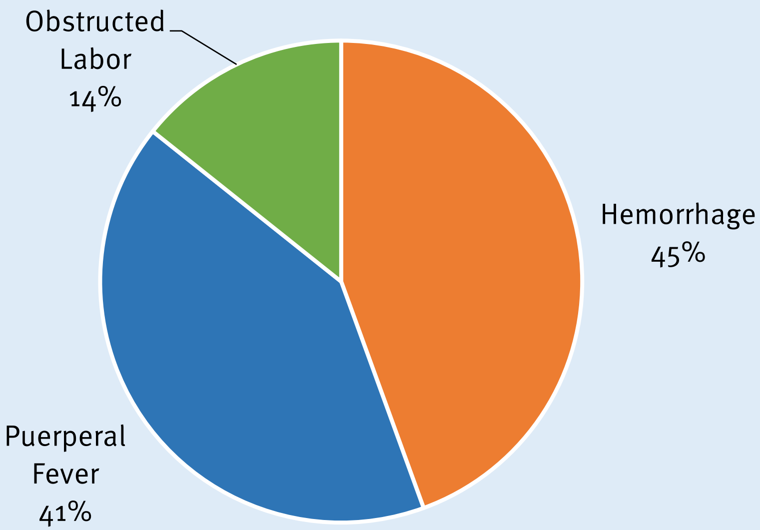
Records from An American General Practitioner, 1811 to 1838 (1,248 Deliveries), Source: Loudon, p. 288.



England, Early 19th Century

Main Causes of British Maternal Mortality, 1831 - 1843

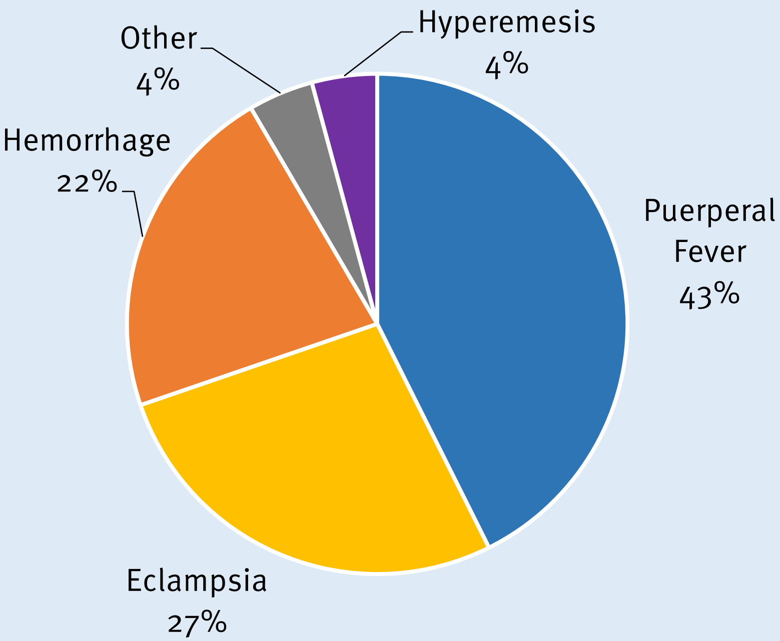
Records from a district of the Royal Maternity Charity (N=126), Source: Loudon, p. 248.



Scotland, Early 20th Century

Main Causes of Scotland Maternal Mortality 1911 - 1915

Source: Loudon, p. 248.



Source: Irvine Loudon, *Death in Childbirth*, Clarendon Press, Oxford, 1992.

The Four Main Historic Causes of Maternal Mortality

1. Puerperal Fever (post-partum septicemia)

- Infection of the uterus and bloodstream following delivery.
- Caused by unwashed hands/instruments of physicians, retained tissue, or unhygienic birth conditions.
- By the 18th–19th c., especially in lying-in hospitals, puerperal fever became the single biggest killer, responsible for up to 40–50% of maternal deaths in some settings.

2. Post-Partum Hemorrhage

- Hemorrhage caused by retained placenta or uterine rupture.
- Often described in early texts as “floodings” or “floods.”
- In the pre-transfusion era, massive bleeding was almost always fatal.

3. Obstructed Labor

- Malpresentations (breech, shoulder dystocia), pelvic deformity (rickets, injuries), or very large babies could trap women in labor for days.
- Before safe Caesarean delivery and forceps many women died from exhaustion, rupture, or infection.

4. Eclampsia / Convulsions

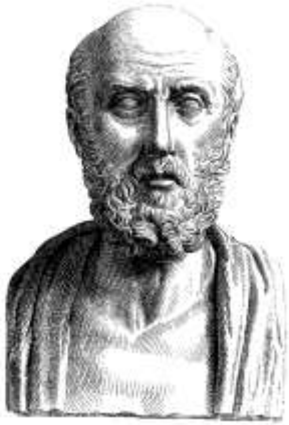
- Hypertensive disease of pregnancy (pre-eclampsia/eclampsia).
- Observed and described since antiquity as sudden seizures in pregnancy or labour.
- Usually fatal without modern interventions.

Chentrey Monument to Anna Greaves who died in childbirth 1819 (Waterperry Church).

Progress in Conquering Puerperal Fever (Ancient Times to 1800)

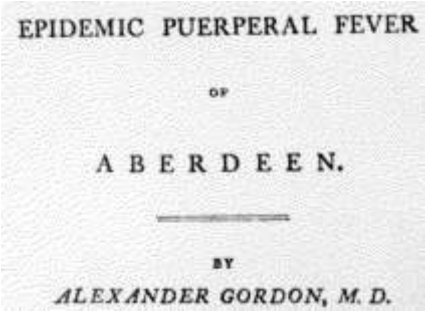
Ancient Times

Hippocrates describes the first known case of puerperal fever.



1651

William Harvey gives a clear description of puerperal fever but has no explanation for it. This followed Jean Morel's 1641 essay on the topic "febre purpurata".


EPIDEMIC PUERPERAL FEVER

OF

ABERDEEN.

BY

ALEXANDER GORDON, M. D.

1795

Alexander Gordon notes that puerperal fever is introduced from a surgeon's finger who has previously touched a "putrid body".

1495

Ortolff von Baierland provides clear advice to avoid introducing unclean hands to a woman's uterus.

1773

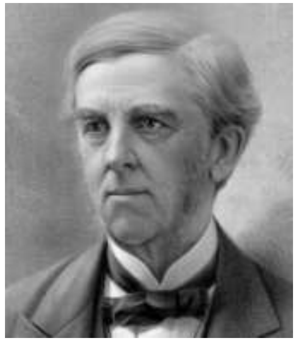
Charles White notes that the vaginal discharge becomes putrid in cases of puerperal fever. John Burton (1751) earlier suggested it is contagious.

Sources: Cutter and Viets, *A Short History of Midwifery*, Philadelphia, W. B. Saunders, 1964 and the author's reading of *Fraubuchlein* by Ortolff von Baierland.

Progress in Conquering Puerperal Fever (1801 to 1936)

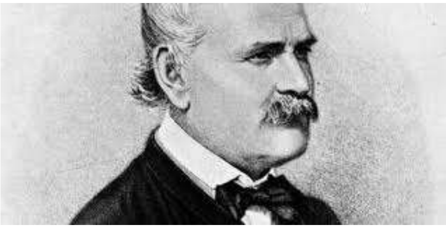
1815-1831

Numerous authors link puerperal fever to unclean conditions in birthing chamber.



1843

Oliver Wendell Holmes forcefully argues that puerperal fever is caused by unclean hands of physicians delivering babies. He is ignored.

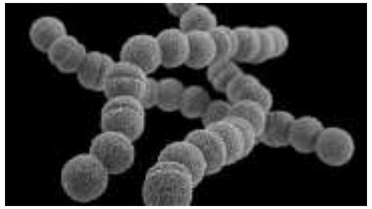


1848

Ignaz Semmelweis proves beyond doubt that puerperal fever is reduced by chlorine disinfection of hands.

1879

Louis Pasteur identifies a bacterium (*streptococcus*) in the blood and uterus of a patient who died of puerperal fever.



1935

Dora Colebrook conclusively typed *Streptococcus* as the main cause of puerperal fever.

1936

Leonard Colebrook (Dora's brother) treats puerperal fever with antibiotics, saving 64 patients in a single year.

Sources: Cutter and Viets, *A Short History of Midwifery*, Philadelphia, W. B. Saunders, 1964 and the author's research.

Explaining the Decline in Maternal Mortality

Going back to the chart on p. 12, there was a fourfold reduction in maternal mortality from 1650 to 1825.

Historians attribute this to the professionalization of midwifery and medical practice.

This period also coincided with rapidly rising living standards and improvements in nutrition which meant that women were stronger going into childbirth.*

The field of obstetrics as a professional discipline emerged in the 1700s and knowledge of how to manage obstructed labor and hemorrhage, in particular, blossomed in this period. In contrast, deaths from puerperal fever remained stubbornly high up until the 1930s.

What is surprising is that society, more or less, understood the cause of puerperal fever by 1795 and knew what to do about it by 1850.

A key message coming out of Semmelweis' work was that Charles White's advice to keep the birthing chamber well ventilated and that obstetricians should wash their hands was not enough.

Semmelweis showed, however, that fingers which previously caused puerperal fever no longer did so when surgeons washed their hands in chlorine.

The specific reason why this mattered was explained quite well in 1879 by Pasteur (and others). Physicians that had performed surgery or dissections had been exposed to germs and antiseptic procedures were well suited to eliminating those germs.

These were not the only voices on the matter. Oliver Wendell Holmes, later Dean of Harvard Medical School, argued the case that puerperal fever was linked to infection quite well.

* See Laslett, Peter. *The World We Have Lost*. Routledge, 1965 and Wrigley, E.A. & Schofield, R.S. *The Population History of England 1541–1871*. Cambridge University Press, 1981.

So, What Went Wrong?

Basically, obstetricians and assorted “accouchers” of the time paid little attention to the publications of Semmelweis, Holmes and others.

Both Semmelweis and Holmes faced violent opposition to their theory that physicians themselves were somehow causing deaths through unsanitary practices.

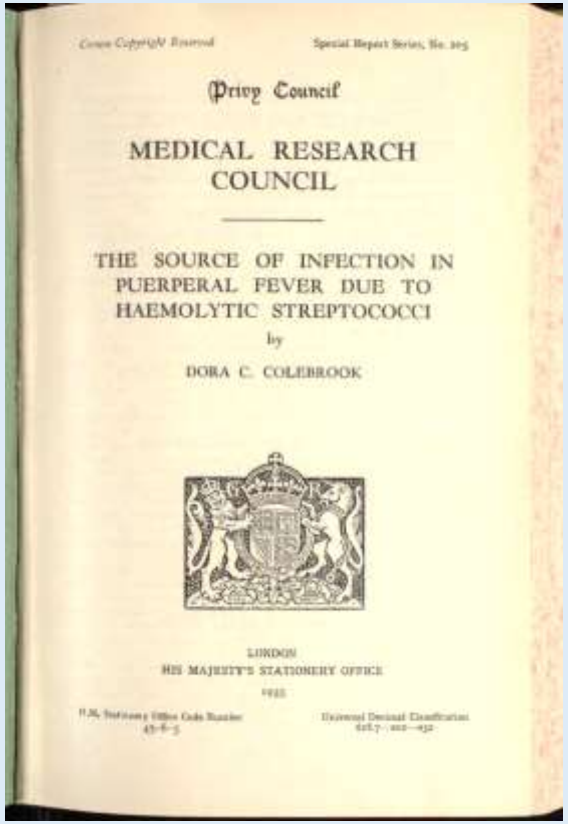
Charles D. Meigs, a leading American obstetrician, famously dismissed Holmes’ claim.* He argued: “Doctors are gentlemen, and gentlemen’s hands are clean.”** In Vienna, Semmelweis confronted entrenched hospital authorities who resisted reforms imposed by a relatively junior figure.

Most 19th-century physicians believed puerperal fever arose from “miasmas,” epidemic constitutions, or humoral imbalance, not direct contagion. This made the idea of physician-borne infection seem implausible.***

* This is the same Charles Meigs who opposed entry of women into the medical profession arguing “The great administrative faculties are not hers…such is not a woman’s province, nature, power or mission. She reigns in the heart; her seat and throne are by the hearthstone. The household altar is her place of worship and service…she has a head almost too small for intellect and just big enough for love.” [Charles D. Meigs, *Females and Their Diseases: A Series of Letters to His Class* (Philadelphia: Lea and Blanchard, 1848)].
** Meigs, Charles D. *A Treatise on the Nature, Signs, and Treatment of Childbed Fevers*. Philadelphia, 1854.
*** Loudon, Irvine. *The Tragedy of Childbed Fever*. Oxford University Press, 2000.

Semmelweis faced such violent opposition to his work on the cause of puerperal fever that today this type of opposition is called a “Semmelweis Reflex.”

Medical Culture of the 19th Century Strongly Opposed New Evidence



Privy Council

MEDICAL RESEARCH COUNCIL

THE SOURCE OF INFECTION IN PUERPERAL FEVER DUE TO HAEMOLYTIC STREPTOCOCCI

by

DORA C. COLEBROOK

LONDON

HIS MAJESTY'S STATIONERY OFFICE

Dora Colebrook’s 1935 paper definitively pointed to streptococci as the source of puerperal fever.

Without germ theory (Pasteur’s work came in the 1860s–70s), there was no conceptual framework for invisible agents carried on the hands. Further, Semmelweis delayed publishing his results until 1861 and when he did, his writing style was combative and alienating.

He became highly distraught and was beaten to death in 1865 in an Austrian mental institution.

Medical culture valued grand theory (miasmas, constitutional disease) over the statistical evidence Semmelweis provided. His insistence on empirical outcomes clashed with medical norms of the time.*

What is so striking to us is how rapidly the insights of Dora and Leonard Colebrook were adopted. Leonard Colebrook showed that sulfonamide antibiotics were highly effective in treating puerperal fever. Within a decade, deaths from this disease plummeted. In other words, preventive practices which required change in physician behavior were slow to be adopted but drug treatment was rapidly adopted.

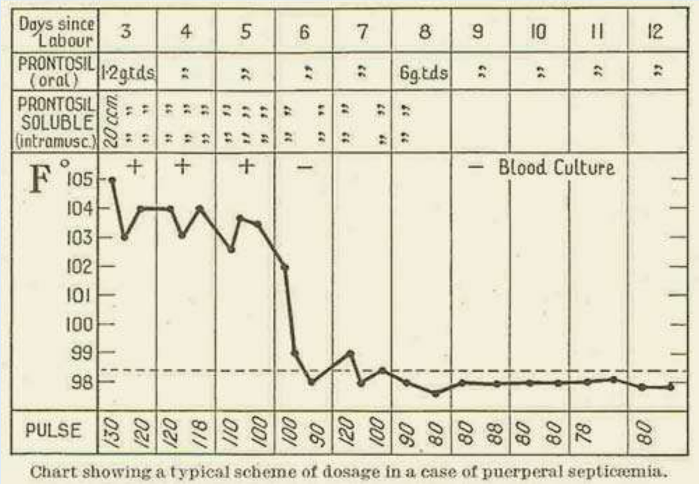


Chart documenting a typical fall in the fever of a patient infected with puerperal sepsis and decline of bacteria in blood cultures following treatment with Prontosil by Leonard Colebrook in 1936

* Porter, Roy, *The Greatest Benefit to Mankind: A Medical History of Humanity*. HarperCollins, 1997.

Change in Medical Practice is Slow

Gerald Weissmann wrote in 1997: "For much of the 19th century many thousands of women died of an 'almost, if not entirely preventable' disease, because public beliefs—in the miasmatic theory, or the benign agency of doctors and midwives—prevailed over the facts of public health."

Lost and Forgotten

Irvine Loudon wrote (p. 68): "Semmelweis' work had **little impact** in Britain, Holmes's essay was still 'largely forgotten' even at the end of the nineteenth century, and Alexander Gordon was almost totally forgotten until the twentieth century."

But physicians ultimately took some notice after the work of Pasteur and the **establishment of germ theory**. Dr. Warren of Portland Maine wrote in *JAMA* in 1905 that before 1870 most students said when they came from the post-mortem room, "What is the point of washing one's hands before making a vaginal examination? Will they no be just as dirty after it?" When those students graduated, they made no attempt to practice antisepsis. Those who graduated between 1870 and 1880 were more conscientious but those who graduated after 1890 and became obstetricians were invariably enthusiasts and "avowed antiseptists".**

Even today, the use of hand hygiene among examining physicians is far from perfect. In 2014, Henderson, Lee and Palmore wrote in *Infection Control Today:*

"Despite the fact that the healthcare industry has had more than a century and a half to assimilate the incontrovertible evidence that hand hygiene prevents the transmission of healthcare-associated infections, the industry has been unable to modify its culture to incorporate the use of hand hygiene as a consistent patient safety practice. Perhaps even more troubling is that 150 years after the publication of Semmelweis' treatise, we continue to encounter a contemporary 'Semmelweis reflex' about the use of hand hygiene in healthcare."***

* Gerald Weissmann, NYU, "Puerperal Priority," *Lancet* 1997, 349, pp. 122-125.
** "Puerperal Infection in Private Practice," *JAMA* 33, 1902, pp. 1008-9 and Loudon pp. 299.
*** David K. Henderson, MD; Laura M. Lee, BSN, MSc; and Tara N. Palmore, MD, "The Contemporary Semmelweis Reflex: History as an Imperfect Educator," *Infection Control Today*, June 1, 2014.

“The frequent sequence of reactions to an important discovery is first a denial of its veracity, then a denigration of its importance, and finally usurpation of credit for it.”

Alexander von Humboldt
*The Great German Explorer (1836)**

* This is a bit of a paraphrase from Alexander von Humboldt, *Examen Critique de L’Histoire de La Géographie du Nouveau Continent*.

Solutions to Obstructed Labor Very Slow to be Adopted

The understanding of why some women are not able to deliver a baby despite days of labor advanced substantially in the mid-1700s and 1800's. Contributors included those who

1. **Came to understand the causes of obstructed labor, including birth process and the role of the pelvis.**

   [William Smellie (1752), Matthias Saxtorph (1771), William Osborn (1783), Carl Caspar Créve (1795), Franz Carl Naegele (1825), and Adolphe Pinard (1892)].

2. **Learned to use forceps to extract a stuck baby**

   [Rueff (1587), the Chamberlens (1600s), Jean Palfin (1708), Edmund Chapman (1733), William Giffard (1734), André Levret (1747), Benjamin Pugh (1754), Jean Astruc (1761), Jean Louis Baudelocque (1775)].**

3. **Learned to resolve obstructed labor with Caesarian**

   [Rousset (1581), James Barlow (1822), Carl Ferdinand von Graefe (1826), Thomas Radford (1865), Porro (1876), Sänger (1882)].

Long Lag Times in Adoption of Innovation

Forceps turned out to be an incredibly helpful way to extract a baby in obstructed labor. While these were first suggested by Jacob Rueff in the 1500s, their use was not widespread until **170 years later**. They were still being improved upon in the early 20th Century.

The **Caesarian Section** is today's answer to obstructed labor and is highly effective in saving both the child and mother. Rousset noted in the 1580's that it could be performed successfully but it was not until the early 1800s that the procedure started to be used.* We did not see widespread use of the C-section until the 1930's to resolve obstructed labor. This was a **gap of almost 400 years**.**

* See Mitchell, Vaughn and Veluri, "The History of the C-Section," *Academic Medicine and Surgery*, Sep 1, 2024.
** Walter Radcliffe, *Milestones in Midwifery and the Secret Instrument: The Birth of Midwifery Forceps,* Norman Publishing, San Francisco, 1989.

Effective Management of Post-Partum Hemorrhage Took Centuries

Perhaps the most feared post-partum complication has been post-partum hemorrhage (the "floodings"). Perhaps the greatest obstetrician of all time, William Smellie, said that this was his most feared condition. Contributors in solving this problem included those who:

1. **Came to understand the causes of post-partum hemorrhage.**

   [Soranus (200AD), Ambroise Paré (1575), Jacob Rueff (1587), Jacques Guillemeau (1609), Marie Anne Boivin (1819)].

2. **Developed obstructive sponges for this problem.**

   [Hildegard of Binden (1150AD), Charles White (1762)]

3. **Learned to resolve placenta previa and ruptured uterus via uterine compression, ligatures, removal of the placenta and other means.**

   [Soranus (200 AD), Pare (1575), Smellie (1752), John Harvie (1767), Jean-Louis Baudelocque (1781), Edward Rigby (1775), Stearns (1822), James Dowling Trask (1855), Crede (1861).]

4. **Learned to use transfusion to save the mother's life.**

   [Rousset (1581), James Blundell (1818), James Barlow (1822), Carl Ferdinand von Graefe (1826), Thomas Radford (1865)].

Long Lag Times in Adoption of Innovation

While the causes of post-partum hemorrhage were somewhat understood in the 16th and 17th centuries it was not until Edward Rigby's work in 1775 that there was a clear understanding of the different causes and potential treatments of post-partum hemorrhage. Progress in dealing with the condition was excruciatingly slow. Some progress was made in the 18th and 19th centuries, but the condition remained a major cause of maternal death throughout this period.

The use of **blood transfusion** to save the mother was not widely adopted until the 1940s despite effective use by Blundell in 1818.*

* See Blundell, J. "Experiments on the Transfusion of Blood by the Syringe." *Medico-Chirurgical Transactions* 9 (1818): 56–92 and P.L. Mollison, *Blood Transfusion in Clinical Medicine*. Blackwell, Oxford (1967).

Progress in Curing Scurvy

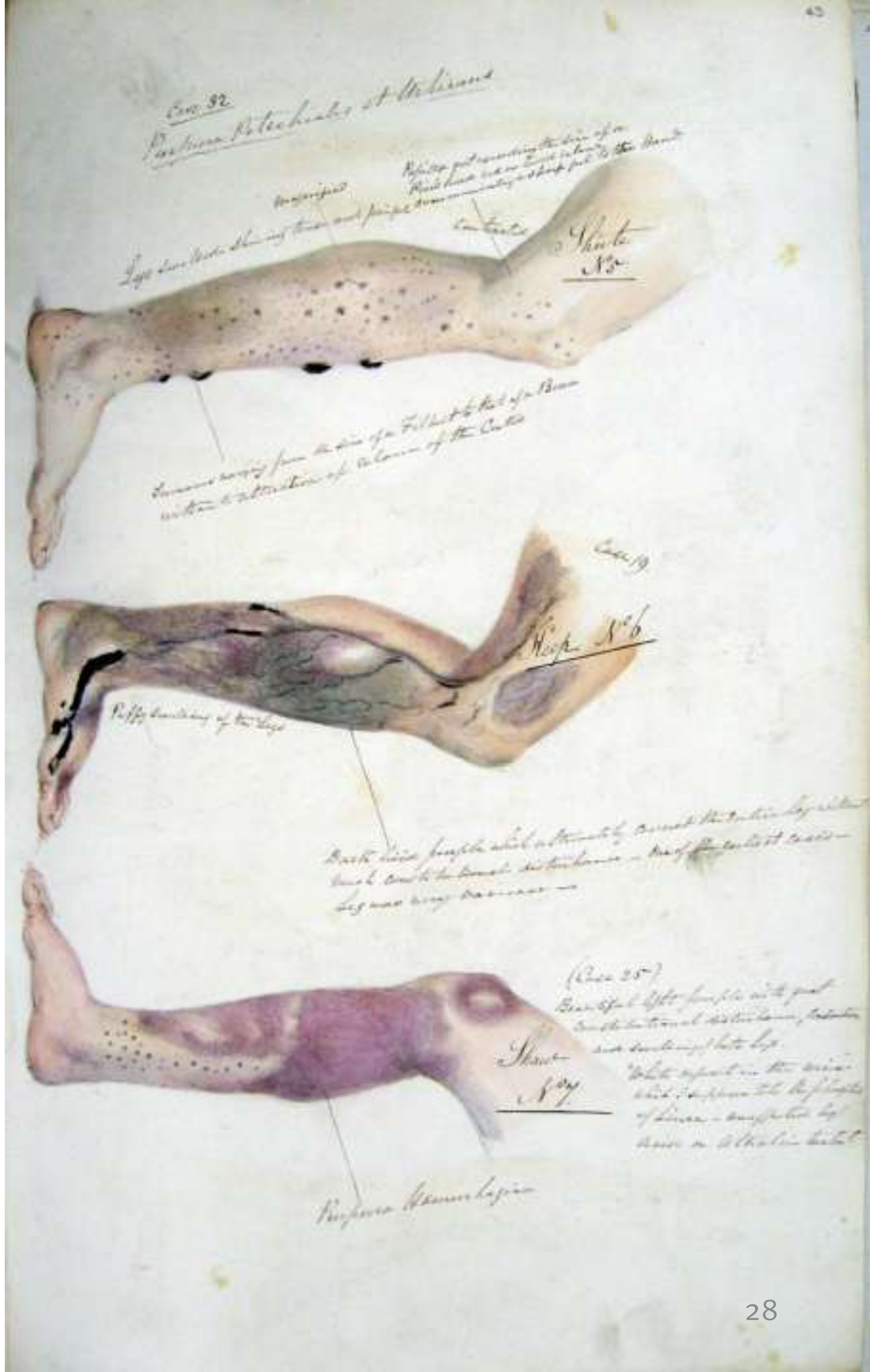
Scurvy is a Disease of Vitamin C Deficiency

Scurvy is a debilitating disease caused by a severe lack of Vitamin C, which is essential for producing collagen, the protein that binds the body's tissues.

On long sea voyages, where fresh food was impossible to preserve, sailors lived on a "preserved" diet: hardtack (a dry biscuit), salted or pickled meat (often years old) and dried grains. This diet was completely devoid of Vitamin C.

The results were catastrophic. Symptoms of scurvy would appear after a month or two at sea:

- **Early stages:** Lethargy, weakness, aching muscles and pale skin.
- **Advanced stages:** Swollen, bleeding gums; teeth falling out; old wounds reopening; severe bruising from internal hemorrhaging and profound weakness.
- **Final stages:** Jaundice, fever, convulsions, and death.



Page from the journal of Henry Walsh Mahon showing the effects of scurvy, from his time aboard HM Convict Ship Barrosa (1841/2)

Over 2 Million Sailors Died of Scurvy

The mortality rates were staggering. It's estimated that scurvy killed more than two million sailors between 1500 and 1800. On some voyages, mortality rates could reach 50% or even 90%.

Anson's famous circumnavigation (1740-44) began with nearly 2,000 men; over 1,400 died, most from scurvy.

Ships would often arrive at their destinations with barely enough men to man the sails, having thrown dozens or hundreds of corpses overboard during the journey.

The medical establishment of the time, operating on the flawed "humoral theory" or miasma theory, was baffled. Proposed causes included: (1) Poor air below decks, (2) Dampness and cold, (3) Laziness (a moral failing), (4) The "taint" of the salted meat and (5) a lack of acid in the body.

Treatments based on these theories, like bloodletting, administering sulfuric acid, or forcing exercise, were not only useless but often accelerated the victim's decline.

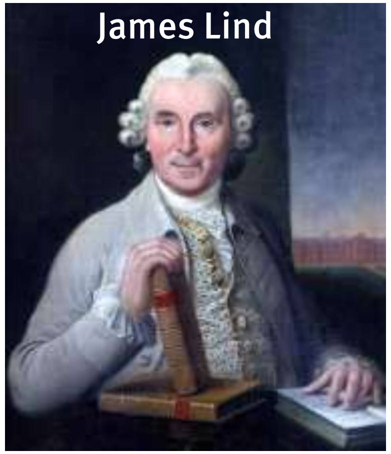
It Took the British Navy 177 Years to Decisively Act to Cure Scurvy Once the Solution Had Been Published in 1617

1497

John Cabot's crew devastated by scurvy on voyage to North America.

1747

James Lind conducts clinical trial and shows that citrus cures scurvy.



1932

Vitamin C synthesized by Szent-Gyorgyi, effectively curing scurvy once and for all.

1601

James Lancaster uses lemon juice on a voyage, and his crew does much better than before. He effectively finds the answer which is published several times including in 1617.

1795

British Admiralty finally includes lemon juice as a standard ration on ships, effectively curing scurvy in the British Navy.

Why the Long Delay?

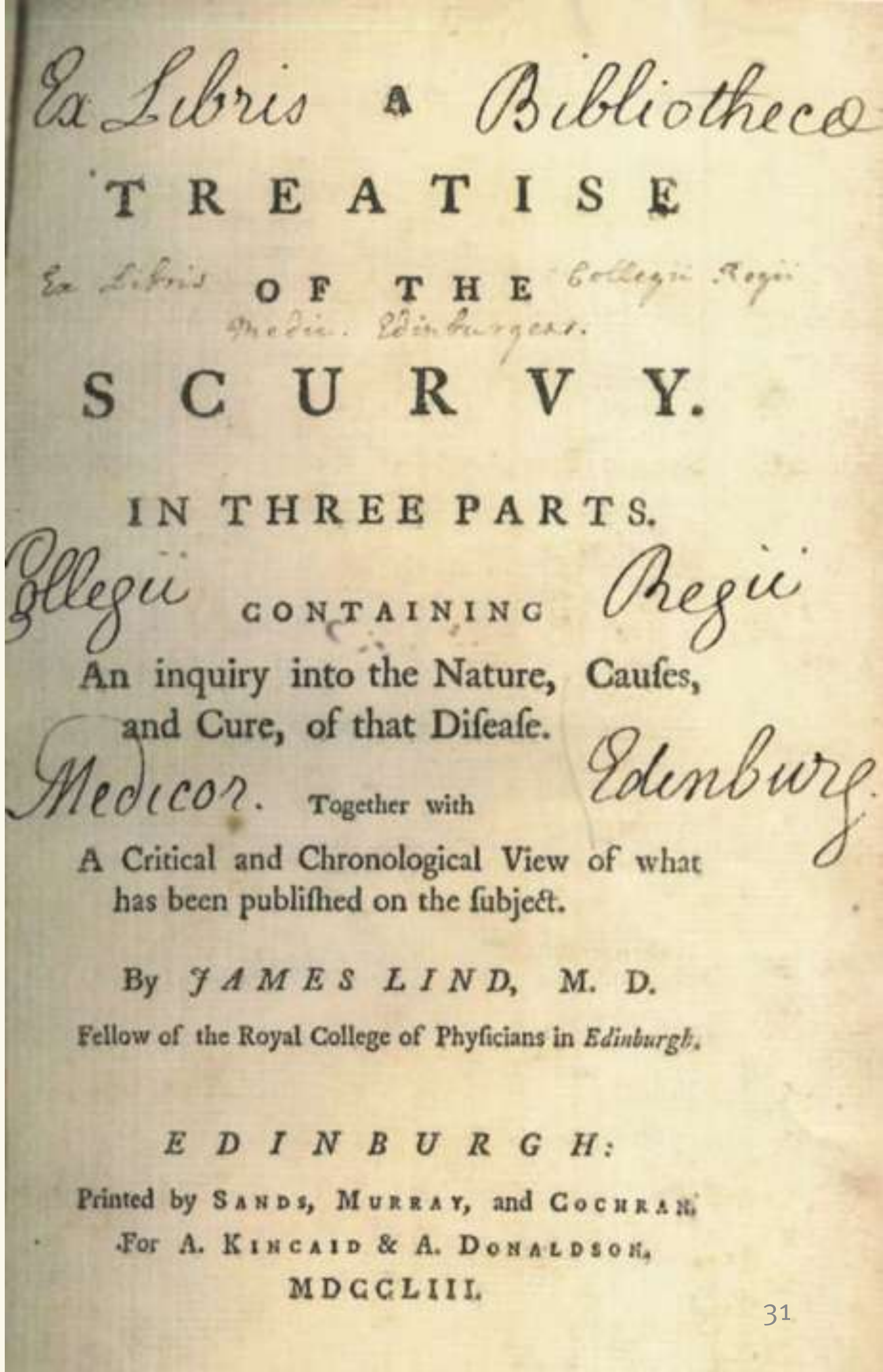
James Lancaster of the East India Company had run an experiment in 1601 on the Red Dragon, a ship rounding the Cape of Good Hope and noted that sailors that took lemon juice did not get scurvy. By 1617, John Woodall, the Surgeon General of the East India Company recommended carrying a good quantity of lemon juice to avoid scurvy in his book *The Surgeon's Mate*.

While the recommendation was in the plain of day it was ignored except by the East India Company, hence devastating losses as the Age of Sail and great navies got underway in the period from 1650 to 1800.

After James Lind's clear rediscovery of the importance of lemon juice in 1747, the British Admiralty did not recommend lemon juice on its ships for another 50 years.

Why?

It turns out that Lind's book on scurvy did not clearly recommend lemon juice. His finding was buried in a long treatise. And an incorrect competing theory (sailors should take sauerkraut) from Captain Cook was much more widely promoted. Cook had the ear of the Admiralty.



A
TREATISE
OF THE
SCURVY.
IN THREE PARTS.
CONTAINING
An inquiry into the Nature, Causes, and Cure, of that Disease.
Together with
A Critical and Chronological View of what has been published on the subject.
By JAMES LIND, M. D.
Fellow of the Royal College of Physicians in *Edinburgh*.
EDINBURGH:
Printed by SANDS, MURRAY, and COCHRAN,
For A. KINCAID & A. DONALDSON,
MDCCLIII.

James Lind: The Man Who Helped To Cure Scurvy

Marcus White, *BBC News*, October 4, 2016 (excerpt)

James Lind is remembered as the man who helped to conquer a killer disease. His reported experiment on board a naval ship in 1747 showed that oranges and lemons were a cure for scurvy. But why did the Royal Navy, which celebrates the tercentenary of Lind's birth on 4 October 2016, take nearly half a century to act on his findings?

For 18th Century sailors, disease during long sea voyages was often more dangerous than enemy action.

One British expedition to raid Spanish holdings in the Pacific Ocean in the 1740s lost 1,300 of an original complement of 2,000 men to illness.

The commander, George Anson, said "almost the whole crew" was afflicted by symptoms including a "luxuriancy of funguous flesh... putrid gums and... the most dreadful terrors".

The explorer Captain Cook recommended malt and sauerkraut, while others swore by "elixir of vitriol" (a dilute solution of sulphuric acid), blood-letting and applying a piece of turf to the patient's mouth to counter the "bad qualities of the sea-air".

Among the array of imaginative remedies were some that proved effective.

Sailors who ate the ship's rats were inadvertently protecting themselves - as the animal synthesizes its own vitamin C. Citrus fruit - another source of vitamin C - had already been suggested as a cure by some.

More than a century later, a learned man fulfilled that wish, with a text which earned him a place in scientific history. Dr Lind's "Treatise of the Scurvy", containing a celebrated review of literature on the disease, appeared in 1753, by which time he was a practising physician in Edinburgh.

He prided himself that he had conquered a condition which "during the last war, proved a more destructive enemy, and cut off more valuable lives, than the united efforts of the French and Spanish arms".

But it was not until 42 years later that the Admiralty first issued an order for the distribution of lemon juice to sailors. Historians still debate why it did not act upon Dr Lind's discovery earlier.

Jane Wickenden, from the Institute of Naval Medicine, said she believed it was partly because Dr Lind's treatise drew no clear conclusions. She said: "The account of the experiment only takes up four pages. The remaining 450 pages deal with other treatments including fresh air and exercise."

Adding to the confusion were rival cures, including the malt and "Sour Kroutt" favoured by Captain Cook.

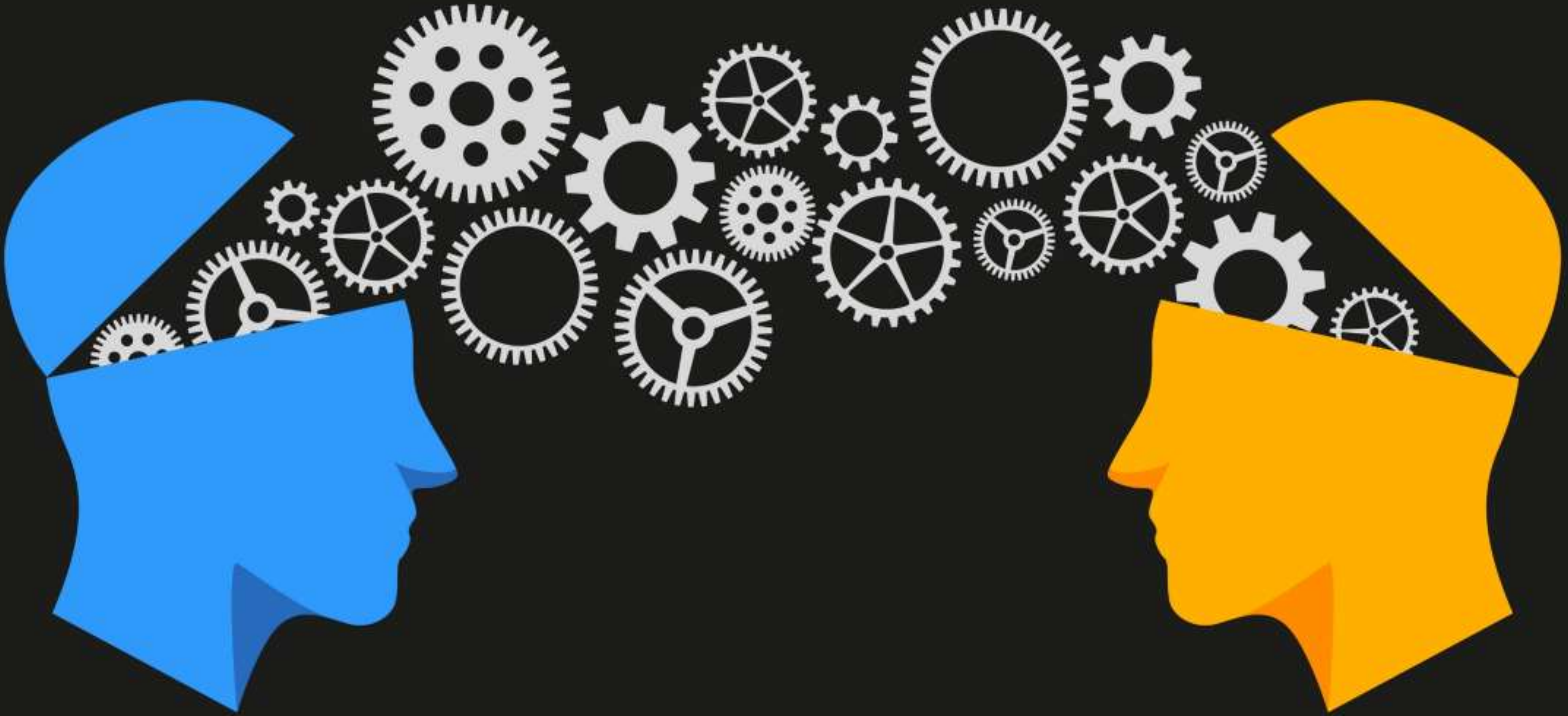
Ms Wickenden added: "Captain Cook was the self-publicist that Lind wasn't. He had travelled around the world and had the ear of the Admiralty."

Source: https://www.bbc.com/news/uk-england-37320399

“There is nothing more difficult to plan, more doubtful of success, nor more dangerous to manage than the creation of a new order of things. . . . Whenever his enemies have occasion to attack the innovator they do so with the passion of partisans, while the others defend him sluggishly so that the innovator and his party alike are vulnerable.”

Niccolo Machiavelli
The Prince (1513)

Knowledge Transmission



Knowledge Transfer is Imperfect

Forgive us for the somewhat long expositions on maternal mortality and scurvy.

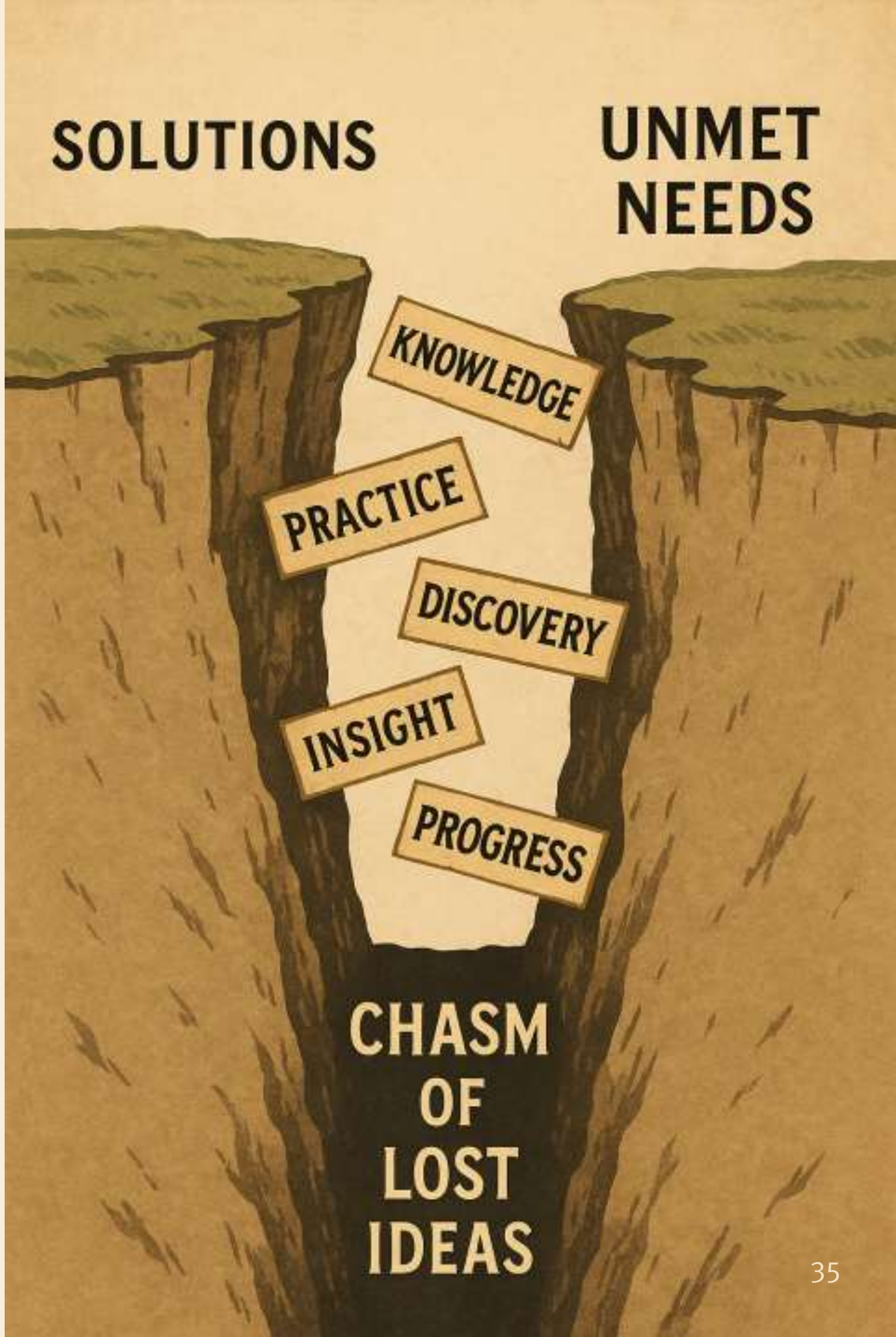
The good news is that most developed societies by now have done a very good job in learning how to reduce maternal mortality. And scurvy is largely gone.

But what is so surprising is how long these journeys took.

In the case of puerperal fever, for example, there was clear advice in Ortolff von Baierland's 1495 book to avoid introducing unclean hands into the mother's uterus.* But it was not until the 1930s that the scourge of death from post-partum septicemia was largely eliminated. Similar stories can be told about dealing with post-partum hemorrhage, obstructed labor and eclampsia.

Routinely, the time from insight to widespread adoption has ranged over 100 years, in some cases more than 400 years. For whatever reason, knowledge transmission, at least in the past, has been far from perfect. Good ideas are routinely ignored, forgotten, lost, opposed by the ignorant, unpublished etc.

* To be fair, von Baierland's book today is found in only four libraries in the world. It is unlikely that the book was widely circulated or read at the time. Further, he did not defend any of his arguments in the book but simply recited his advice in a litany of comments on what problems might happen after birth and what one might do to solve the problems.



The Pace of Translation of Science is Routinely Slow

Ian D. Graham et.al., "Lost in Knowledge Translation: Time for a Map?," *The Journal of Continuing Education in Health Professions*, 2006, pp. 13-24.

Despite the considerable resources devoted to health sciences research, a consistent finding from the literature is that the transfer of research findings into practice is often a slow and haphazard process. This means that patients are denied treatment of proven benefit because the time it takes for research to become incorporated into practice is unacceptably long.

For example, researchers from the United States and the Netherlands have estimated that 30% to 45% of patients are not receiving care according to scientific evidence and that 20% to 25% of the care provided is not needed or is potentially harmful.

Similarly, it is estimated that cancer outcomes could be improved by 30% with optimum application of what is currently known and that at least a 10% reduction in cancer mortality could be achieved in the United States through widespread use of available state-of-the-art therapies.

Key Obstacles in Knowledge Transmission

	Failure to Preserve Insight	Failure to Communicate Insight	Failure to Propagate Insight	Failure to Accept Insight
PROBLEM	A medical insight got lost or forgotten along the way. For example, libraries or printed matter may have been lost or destroyed.	An insight gets made as a throwaway comment in a long book and is not picked up by others. There is a "lost in translation" factor at play.	It's one thing to have an insight. It's another to transmit knowledge. Knowledge needs to be communicated in a persuasive way and taught to others, often with repetition.	Incumbents committed to the old way of thinking block acceptance of a new idea. This may be motivated by profit, ego, inertia or stupidity.
EXAMPLE	Soranus' key insights on derisking birth were lost in the fall of Rome. While some of his ideas were preserved up to 600AD in the writings of Aetius and Muscio they were then completely lost to the Western world until a Muscio manuscript was rediscovered in the 1600s by Conrad Gesner.	Semmelweis did not publish his findings for 13 years and a colleague had to report them in an obscure medical journal. When he finally published his findings, his book was poorly written and alienating to the reader.	The Chamberlens invented a pair of obstetric forceps that were unequalled for centuries. They kept the idea of forceps as a means to extract babies in the case of obstructed labor a secret. The value of forceps as a solution was delayed for well over a 100 years.	The British Admiralty was notoriously "clubby" and staffed by the British nobility – who were unlikely to rely on empirical evidence and adopt a good innovative idea. Thus, a cure for scurvy was successfully diverted for over 150 years.*

* The British Admiralty is far from the only military organization that is set in its ways. A report on the U.S. Department of Defense by the Defense Innovation Board in 2024 wrote: "In examining the Department's current state, a series of systemic challenges became evident, collectively fostering an environment of risk aversion and complacency."

Does the Medium of Communication Matter? Historical Means of Transmitting Knowledge

Oral Communication
(Prehistoric Times)

Kronick (1984): "Before [writing] society depended entirely on oral transmission from one generation to another to retain its historical and technical heritage. In spite of this limitation, elaborate belief systems and an impressive technology developed involving agriculture, astronomy, the ability to manufacture both domestic utensils and instruments of war."

Writing & Manuscripts
(5000 Years Ago)

Kronick (1984): "The invention of writing is probably the most important technological advance which occurred in communication. It very likely has not existed for more than five or six thousand years, and before this society depended entirely on oral transmission from one generation to another to retain its historical and technical heritage."

Printing
(1450 to Present)

Kronick (1984): "The introduction of printing, however, ultimately radically transformed the nature of scholarship and the transmission of information. Texts available in identical copies could be disseminated more widely than had been possible with manuscript books."

Web / Electronic Media
(1984 to Present)

While electronic copies of books and journal articles had been available for transmission via email, listservs and CD-Roms, it was the introduction of the Worldwide Web in 1984 that made it possible for scientists to more easily share and access information in a revolutionarily easy format.

Source: David A. Kronick, Literature of the Life Sciences: Historical Background, *Bulletin of the New York Academy of Sciences*, November 1984.

The Medium Clearly Mattered - Although Adoption Lags Did Not Disappear With Print

Before print (pre-1450s), medical discoveries were largely transmitted via handwritten manuscripts. These were:

- **Scarce**: a single copy could take months to produce.
- **Localized**: surviving primarily in monasteries, universities, or a limited circle of practitioners.
- **Vulnerable**: copying errors, loss in war or fire, and lack of standardization limited dissemination.

After print (post-1450s), knowledge circulation accelerated:

- **Rapid multiplication**: thousands of copies possible.
- **Standardization**: Galen, Avicenna, Hippocrates, and their commentators became widely available in consistent editions.
- **New genres**: printed midwifery manuals, anatomical atlases, and plague treatises reached practitioners beyond the university.
- Yet—even in the print era—implementation lagged. For example, William Harvey's *De motu cordis* (1628) was in print and quickly known across Europe, but it took many decades before physicians abandoned Galenic theories of the heart in daily practice.

Studies and Publications on the Medium of Communication

Elizabeth Eisenstein, *The Printing Press as an Agent of Change* (1979) argues that the shift from manuscript to print culture was a revolutionary transformation that fundamentally reshaped the Renaissance, the Reformation, and the Scientific Revolution by standardizing knowledge, increasing its accessibility, and fostering new forms of intellectual and social organization. She also notes that inertia in medicine remained strong.

Harold Cook, *Matters of Exchange* (2007) shows how print enabled the spread of medical-botanical knowledge in the early modern Dutch empire, but implementation still depended on trade networks and institutional adoption.

More recently, studies on **knowledge translation in biomedicine** (e.g., Balas & Boren, *Managing Clinical Knowledge for Health Care Improvement*, 2000) show modern publication (journals, databases) shortens but does not eliminate lag—still averaging 17 years for new knowledge to reach practice.

The History of Information: The Web Has Revolutionized Media

Jeremy Norman, *Essay on The History of Information*, 2021 (excerpt)

The existence of diverse electronic media, including telephone, radio, and television, and the way that they have converged on the Internet, has complicated any comparison between the much simpler revolution from manuscript copying to print in the 15th century and that of our time, yet from our vantage point in 2020 it is evident that many of the same factors are in play.

From its slow beginning in 1945 digital information became increasingly dominant, all the while traditional physical print media continue to play significant roles, much the way manuscript copying persisted to a gradually diminishing extent after printing predominated at the end of the 15th century.

If we want to associate the two media revolutions with individual inventors, we might want to compare the achievements of Johannes Gutenberg with those of Tim Berners-Lee, inventor of the World Wide Web.

Like Gutenberg, in his invention of the World Wide Web Berners-Lee built on a complex set of existing technologies underpinning the Internet that had been invented in the roughly fifty years since the invention of the first electronic digital computers; however, unlike Gutenberg's invention that is known only from extant examples of his earliest printing, Berners-Lee's invention is completely documented between 1989 and 1991. As we look back over the roughly thirty years since Berners-Lee invented the World Wide Web we may recognize that the World Wide Web was the key invention that integrated the different digital technologies comprising the Internet, catalyzing the digital information revolution around the world, making computer technology accessible to billions of people.

Are we living through a media revolution like the one that occurred in the mid-15th century? If we define a cultural revolution as a paradigm shift rather than necessarily a complete replacement of one medium for another, the brief answer is clearly yes. That being the case, how much does the question matter?

Having immersed myself in the details involved in answering this question for twenty years, I have also concluded that understanding the complex multi-faceted development of information and media that lead to the present matters more than attempting to compare complex developments five centuries apart.

Source: https://www.historyofinformation.com/story.php?t=Essays#essay-media-revo

Literature on Spread of Innovations

It turns out that there are over 5,000 papers that address the question of how knowledge of innovations diffuses throughout societies.

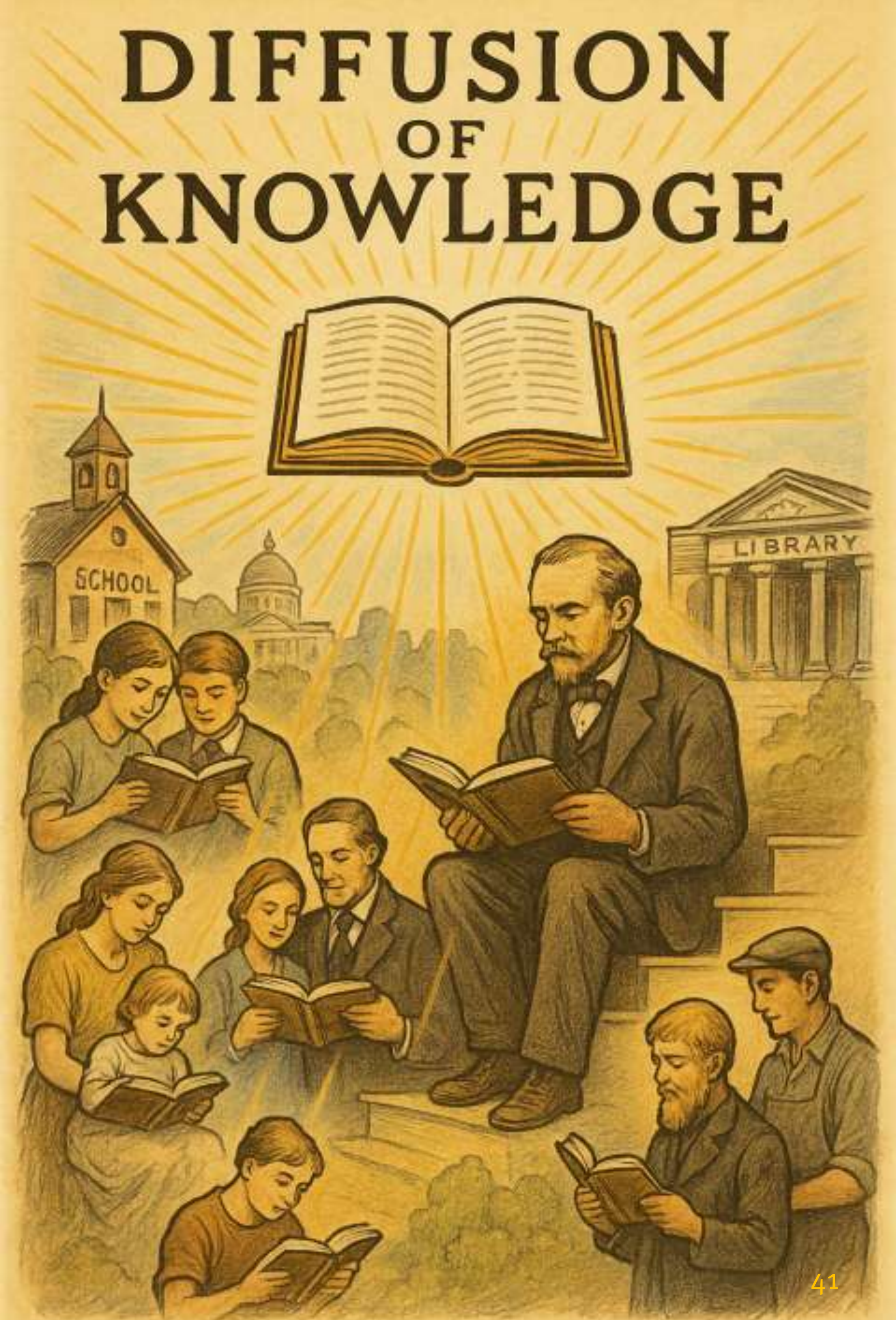
This literature also addresses the puzzling history of good ideas get lost or completely ignored.

Our focus here on the diffusion of knowledge of medical innovations is but one branch of this very large literature that has been largely developed for six decades by sociologists, anthropologists and information theorists.

There is a separate large literature on information theory that discusses how information itself is structured (e.g., the work of Shannon and Kolmogorov) that is less relevant for our discussion.

A quick read through of the literature on medical knowledge diffusion reveals that the experience in maternal health and scurvy is not so unusual. A *repeated pattern* is that a good idea is uncovered but does not get into the standard of care for a long time.

There is an interesting literature that argues that some ideas like Semmelweis' were *premature*. That is, the world wasn't ready for the new ideas.* This is an odd literature that argues for a certain Hegelian historical determinism. We think this literature misses deeper factors that underly innovation adoption.



* See Ernest B. Hook (editor), *Prematurity in Scientific Discovery: On Resistance and Neglect*, University of California Press, 2002 and Stent GS. Prematurity and uniqueness in scientific discovery. *Scientific American*, December 1972, pp. 84-93.

Everett Rogers' Classic Work

Everett Rogers of Stanford University wrote the classic book on the diffusion of innovation in 1962.* He revised the book over several decades as the literature on innovation diffusion grew.

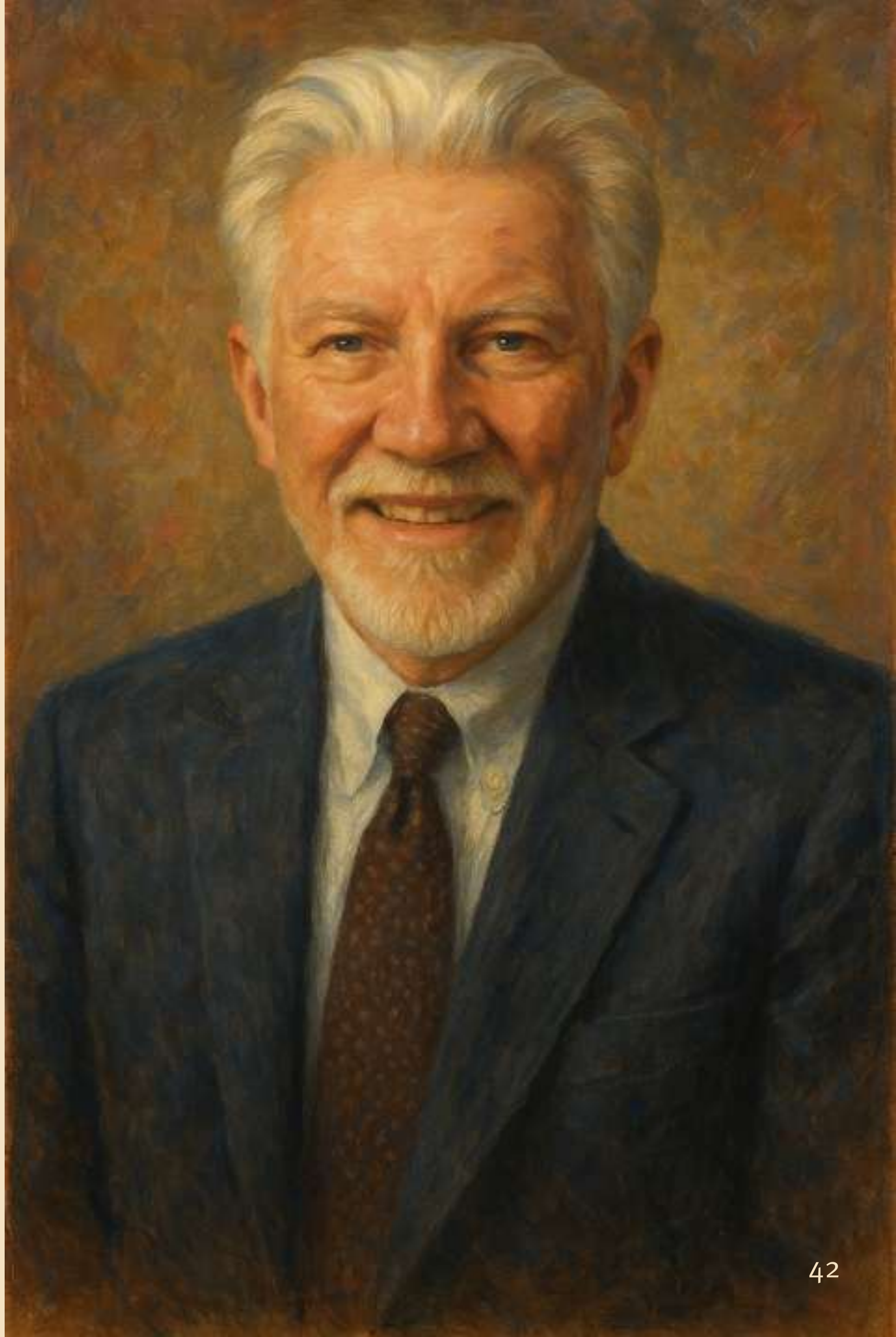
He gives numerous case studies of successful and failed innovation diffusion.

Rogers argued that five key attributes of innovations (relative advantage, compatibility, complexity, trialability, observability) determine their successful adoption (or not).

One central reason is the way individuals perceive the **relative advantage** of the innovation compared to existing solutions. If an innovation does not clearly demonstrate that it is better—whether in terms of efficiency, cost, status, or convenience—adoption tends to lag.

Another factor is **compatibility** with existing values, experiences, and needs. Innovations that clash with cultural beliefs, social norms, or established habits are less likely to be adopted, regardless of their objective merits. For example, medical or agricultural innovations often meet resistance if they contradict local traditions.

Everett M. Rogers, *Diffusion of Innovations*, The Free Press: New York, 1962.



Rogers' Five Attribute Model

ATTRIBUTES OF INNOVATIONS

RELATIVE ADVANTAGE

The degree to which an innovation is seen as better than what preceded it.

COMPATIBILITY

The degree to which an innovation is consistent with existing values and experiences.

COMPLEXITY

The degree to which an innovation is perceived as difficult to understand and apply.

TRIALABILITY

The degree to which an innovation can be tried out on a limited basis to begin.

OBSERVABILITY

The degree to which the results of an innovation are visible to others.

Everett Rogers: Innovations are Adopted Within Social Systems

Rogers, a sociologist by profession, argues that innovation diffusion is the process by which an innovation is communicated through certain channels over time among the members of a social system. Given that decisions are not authoritative or collective, each member of the social system faces his/her own innovation-decision that follows a 5-step process:

1) **Knowledge** – person becomes aware of an innovation and has some idea of how it functions,

2) **Persuasion** – person forms a favorable or unfavorable attitude toward the innovation,

3) **Decision** – person engages in activities that lead to a choice to adopt or reject the innovation,

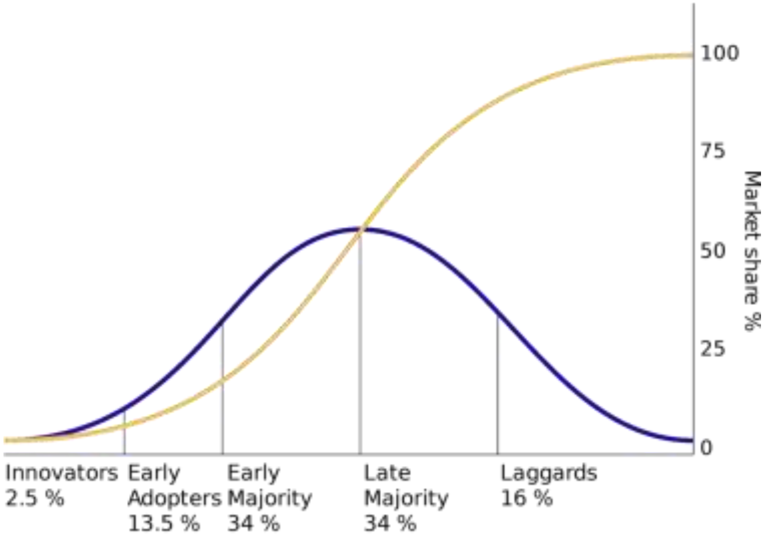
4) **Implementation** – person puts an innovation into use,

5) **Confirmation** – person evaluates the results of an innovation-decision already made.

The most striking feature of diffusion theory is that, for most members of a social system, the innovation-decision depends heavily on the innovation-decisions of the other members of the system.

Everett Rogers Describes an S-Shaped Adoption Curve for Innovation

The Innovation-Adoption Curve is a graphical representation of a model created by Everett Rogers as a method of explaining how, why, and the rate at which an innovation spreads through a population or social system.



The Innovation-Adoption Curve, as described in Roger's work takes a very different approach to describing change than most other behavioral theories. Instead of focusing on persuading individuals to change, it describes change as being caused by the evolution or "reinvention" of products and services, so they become a better fit for the needs of individuals and groups.

This is because the **adoption of new products or behaviors involves the management of risk and uncertainty.** It's usually only people we personally know and trust – and who we know have successfully adopted the innovation themselves – who can give us credible reassurances that our attempts to change won't result in embarrassment, humiliation, financial loss or wasted time.

Early adopters are the exception to this rule. They are on the lookout for advantages and tend to see the risks as low because they are comfortable with high-risk propositions and unproven innovations. Many times, an early adopter will embrace an innovation on the basis of nothing more than a well-worded scientific paper or news article. The rest of the population, however, see high risks in change, and therefore requires assurance from trusted peers that an innovation is adoptable and provides predictable benefits.

As an innovation spreads from early adopters to the mainstream early majority, face-to-face communication becomes more essential to the decision to adopt. Rogers notes that by 2003 there had been eight randomized controlled trials – the gold standard in evaluation – all of which demonstrated the ability of **opinion leaders in peer groups to produce behavioral change.**

Financial Incentives Also Influence Invention Adoption

It is well understood in today's biotech industry that financial incentives are incredibly important in physician decisions to use new drugs on their patients.

Those familiar with the market for AMD drugs, for example, are well aware that expensive buy and bill pharmaceuticals do quite well even though a generic of bevacizumab is far less expensive and effective. The physician gets a payment that is proportional to price and tends to choose the drug that will maximize his or her payment.

Notions of financial incentive are perhaps more comfortable to economists than sociologists like Everett Rogers but they do seem to matter.

Eric Ford (2019) noted, for example, that adoption of electronic healthcare record systems by hospitals was quite slow because of adverse incentives: "Hospitals resisted [EHR] adoption because the systems were a significant added expense, their efficacy as a safety tool was unclear, and the paying patient was largely indifferent. The impetus to adopt an EHR had a low internal motivation factor. Internal influences or desires to adopt are far more powerful in accelerating adoption."*

SOCCER BALL PRODUCTION IN PAKISTAN

Columbia Business School Report: In a report entitled "Organizational Barriers to Technology Adoption: Evidence from Soccer-Ball Producers in Pakistan," published in the *Quarterly Journal of Economics*, the researchers acquainted cutters and printers – workers critical to the process of producing soccer balls – with a way to save material, reducing overall production costs by one percent, a significant amount in a business with low profit margins. It turned out that while it was less wasteful, the production method was slower, which meant less income for workers, who were typically paid piece rates. To alleviate this "misalignment of incentives," a second experiment was conducted among the treatment firms that had received the equipment and details of the production method in the first experiment. Two workers at half of those companies (again chosen at random) were offered bonuses equal to about a month's pay if they could show the owner that they were proficient at using the equipment. Subsequently, after six months, five of them began to use the new technology while none of the companies in the control group, where the incentive pay plan was not offered, did."

* Editor(s): Ford, Eric W. PhD. Paint the Stick Orange: Incentives, Rewards, and the Innovation Imperative. *Journal of Healthcare Management* 64(6):p 349-351, November-December 2019.

Anesthesia For Labor Pain Was Adopted Rapidly

Obstetricians were strongly financially incentivized to use anesthesia for labor pain after its discovery in 1847. This was not the case with handwashing to avoid puerperal fever which was discovered in the same year.

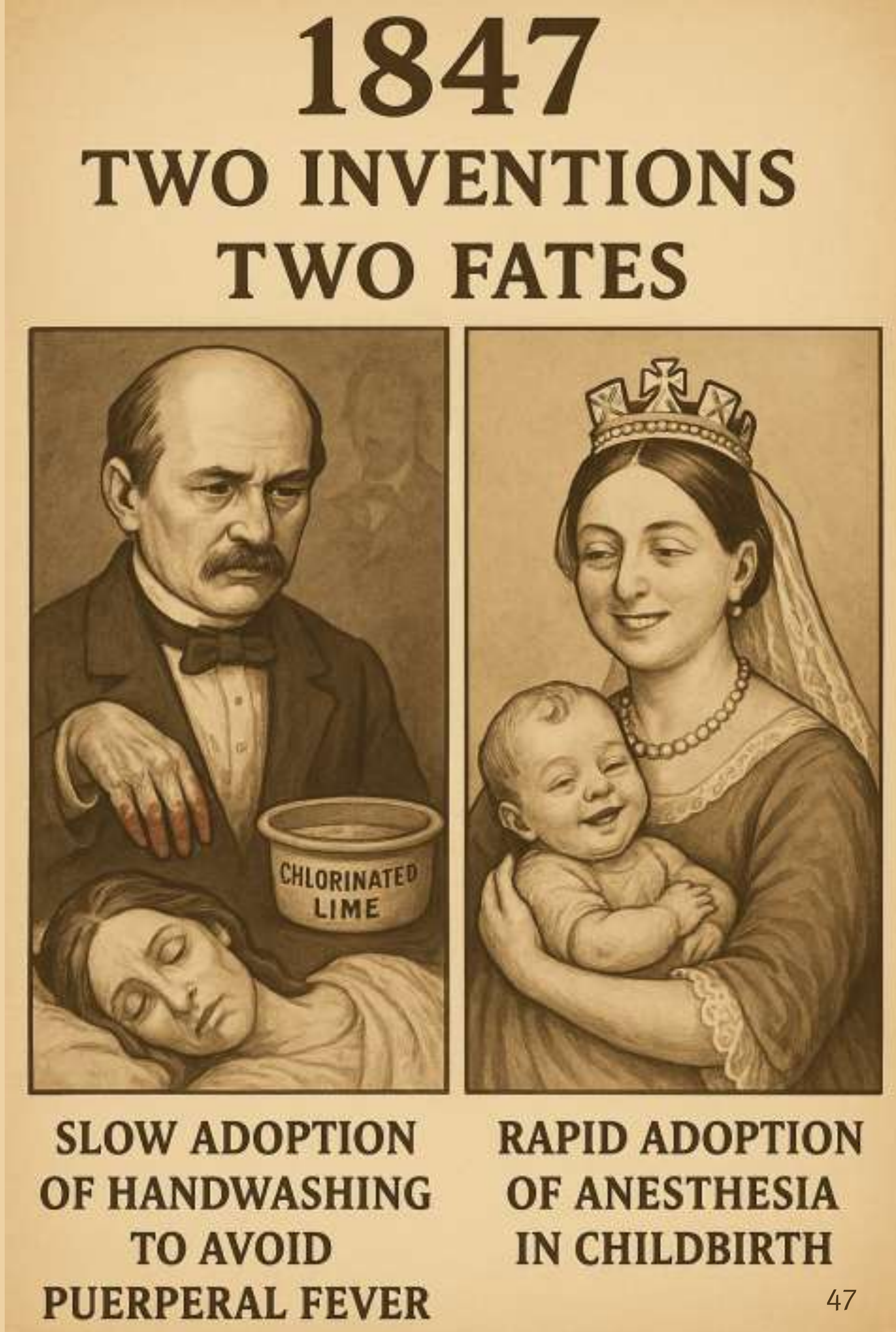
In the mid-19th century, obstetric anesthesia and antiseptic techniques emerged nearly side by side but followed very different trajectories of acceptance.

In 1847 James Young Simpson introduced chloroform in obstetrics. Like Semmelweis, Simpson was an unknown player. Although there were some professional hesitations, the visible relief of labor pain made anesthesia appealing both to women and their physicians.

Anesthesia during childbirth was broadly adopted immediately. Importantly, wealthy women clients were willing to pay extra money to receive anesthesia when delivering a baby and so it was profitable for physicians to use the innovation with their clients.

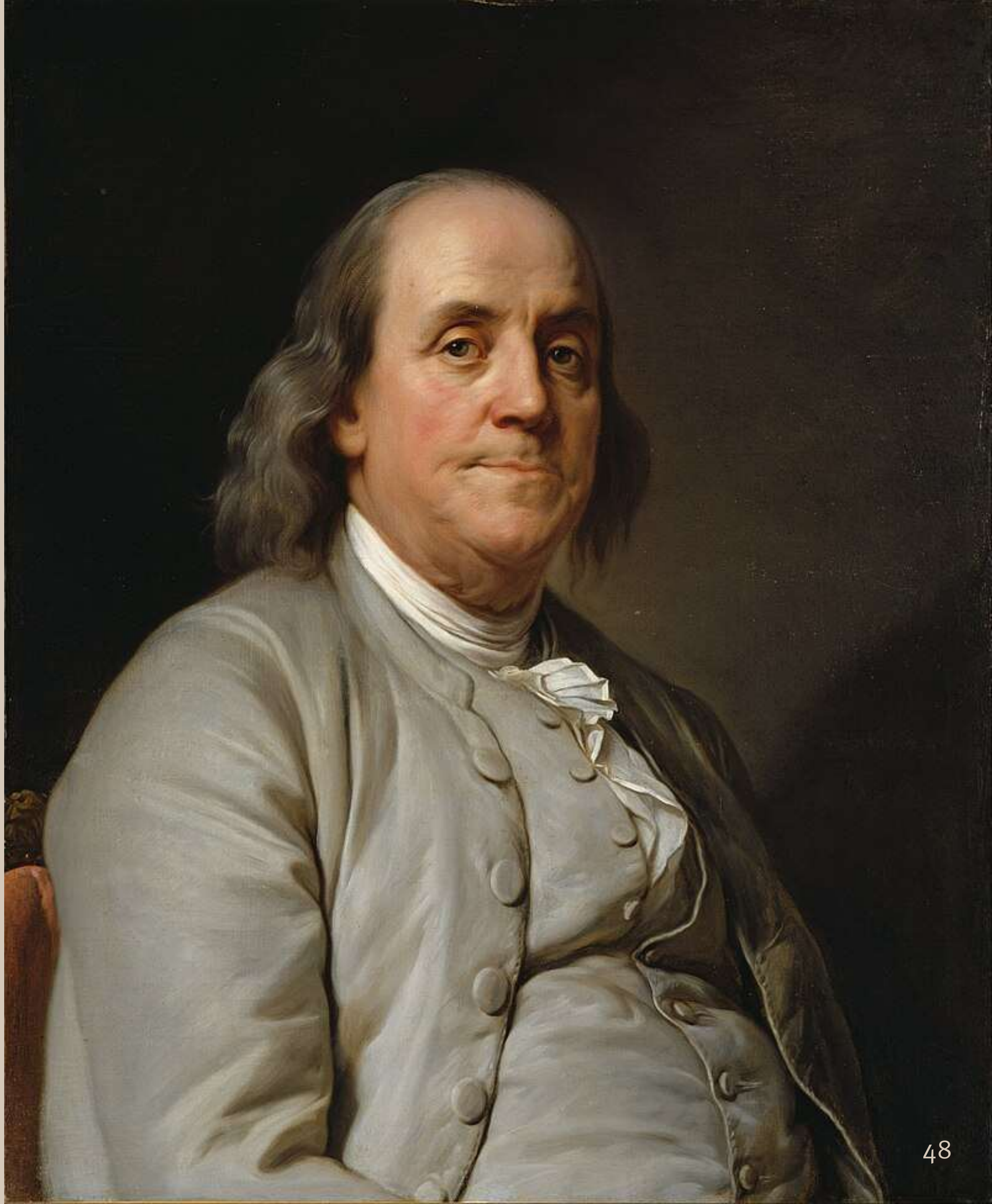
Queen Victoria accepted chloroform during childbirth in 1853, giving the practice a social legitimacy that spurred even broader adoption. Within a decade, use of anesthetics in labor was widespread.

By contrast, Ignaz Semmelweis in Vienna demonstrated in the late 1840's that handwashing with chlorinated lime dramatically cut maternal deaths from puerperal fever. The data were clear, with mortality dropping from around 10% to under 2%. Yet the idea met stiff resistance. Obsetricians bristled at the implication that they themselves were vectors of infection, and without germ theory the rationale behind the practice seemed unconvincing. Importantly, using chlorinated lime required extra work and didn't lead to extra profits for physicians – hence it was convenient for many physicians to ignore the advice.



To get the bad customs of a country changed and new ones, though better, introduced, it is necessary first to remove the prejudices of the people, enlighten their ignorance, and convince them that their interests will be promoted by the proposed changes; and this is not the work of a day.

Benjamin Franklin (1795)

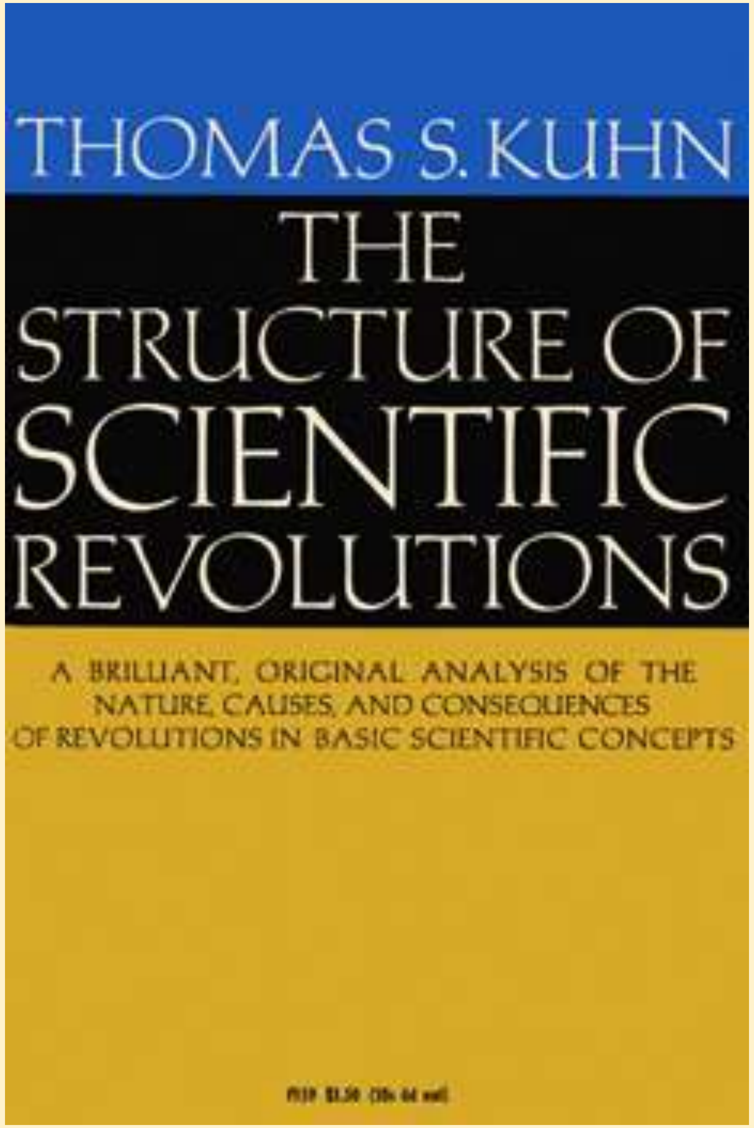


Thomas Kuhn’s 1962 Work

Thomas S. Kuhn’s *The Structure of Scientific Revolutions* (1962) argues that science does not progress in a simple, cumulative way but rather through cycles of stability and upheaval. In Kuhn’s view, “normal science” operates within a **paradigm**—a shared framework of theories, methods, and assumptions—that guides research and problem-solving.

Over time, anomalies accumulate that the paradigm cannot explain, leading to a **crisis** and eventually a **scientific revolution**, in which the old paradigm is replaced by a new, incommensurable one (for example, the shift from Newtonian to Einsteinian physics). Kuhn emphasized that these paradigm shifts are not purely logical but also social and historical, involving persuasion, generational change, and shifts in worldview. His book fundamentally challenged the idea of science as a purely objective, linear accumulation of truth, noting its discontinuous, community-driven, and sometimes subjective character.

Everett and Kuhn both emphasize that change in scientific views requires persuasion of a community of scientists and that this might not be easy.



Graham's KTA Framework Sees Education and Persuasion as the Key to Getting Physicians to Adopt New Approaches to Medicine

Knowledge to Action (KTA) Calls for Physician Education

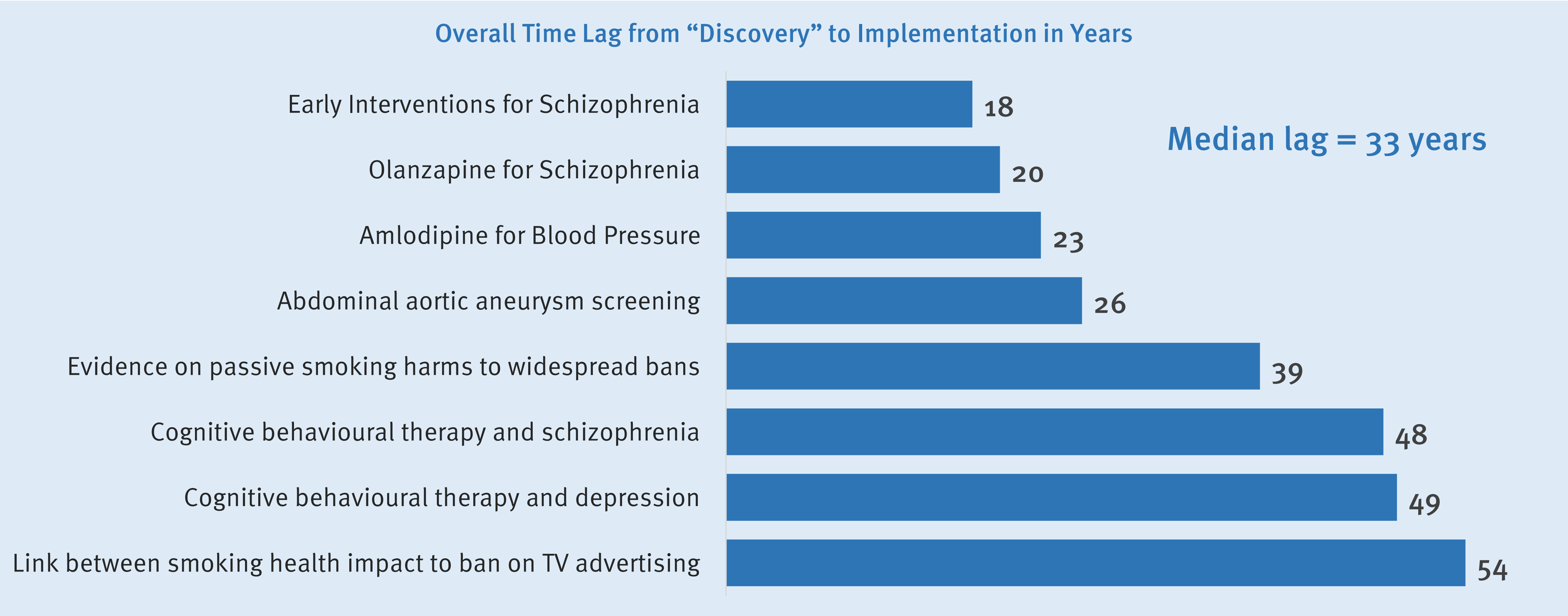
1. Knowledge translation is about turning knowledge into action and encompasses the processes of both knowledge creation and knowledge application.
2. Knowledge translation subsumes and builds on continuing education and continuing professional development.
3. Health care professionals need to learn about planned-action theories and frameworks so as to be able to understand and influence change in their practice settings.
4. Continuing education should be based on the best available knowledge, whether in the form of knowledge tools (e.g., practice guidelines), knowledge syntheses, or primary knowledge inquiries.
5. Continuing education should apply strategies shown to be effective at transferring knowledge.

Structured efforts to educate physicians make sense. Cabana et.al. (1999) studied the literature on physician non-adherence to practice guidelines and found that the leading explanation was physician's lack of awareness, rather than disagreement with guidelines.

Sources: Ian D. Graham et.al., "Lost in Knowledge Translation: Time for a Map?," *The Journal of Continuing Education in Health Professions*, 2006, pp. 13-24 and Michael Cabana et.al, Why Don't Physicians Follow Clinical Practice Guidelines?," *JAMA*, October 20, 1999, p. 1458-67.

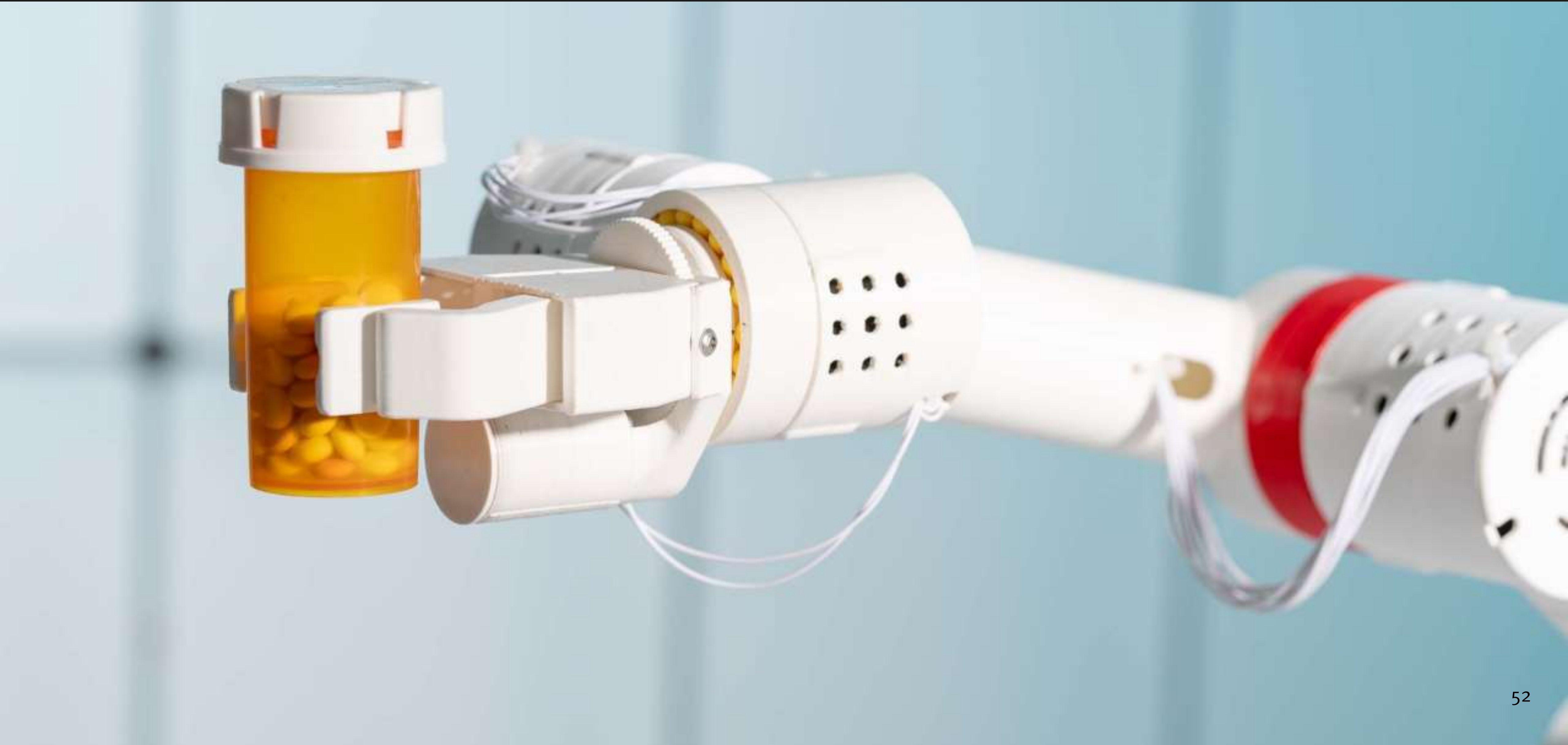
Despite Current Technologies the Median "Bench to Bedside" Gap is Over 30 Years in Healthcare

Hanney et.al. Study (2015)



Source: Hanney SR, Castle-Clarke S, Grant J, Guthrie S, Henshall C, Mestre-Ferrandiz J, Pistollato M, Pollitt A, Sussex J, Wooding S. How long does biomedical research take? Studying the time taken between biomedical and health research and its translation into products, policy, and practice. *Health Res Policy Syst*. 2015 Jan 1;13:1.

How is AI Changing Medicine?



Today We Have a New Medium of Communication: The LLM

While Large Language Models (LLM’s) may not be quite as important as the printing press, the introduction of machine intelligence via the LLM and associated tools is bringing a profound change in how societies process and access information. Unlike all past human communication media, the LLM differs in that the *machine* rather than humans is sharing insights on a topic of interest.

The Large Language Model was introduced very recently and now makes it possible for the computer to generate summaries of the literature and to generate new literature in ways that were previously not possible with the Web and search engines. This creates the seeming impression of thinking machines.

LLM’s read and learn from scientific literature using a process called pretraining and then apply that knowledge through contextual pattern recognition, not conscious reasoning.

LLM’s are not just tools for communicating scientific ideas—they’re becoming collaborators in generating, interpreting, and spreading them. That **does** change the structure and flow of scientific discourse.

LLM’s don’t do science—they simulate the style of scientific reasoning. That means they can generate hypotheses, like suggesting a new drug target based on known pathways. They might propose research directions based on reasoning through the existing literature.

LLM's Make It Easier to Access the Literature and Translate Science to Solutions for Patients

1. **Pinpoint Knowledge Access:** LLM's can find "needles in haystacks" and provide exactly the right knowledge when a specific medical problem arises.
2. **Democratization of Knowledge:** LLMs can explain complex medical concepts to non-specialists, students, or cross-disciplinary collaborators in plain language.
3. **Language Barrier Reduction:** Physicians from non-English-speaking backgrounds can better participate in global science, draft papers in English, or understand foreign-language research more easily.
4. **Literature Review:** LLMs can rapidly synthesize large volumes of medical literature into summaries, highlight trends, or generate research questions.
5. **Peer Review Assistance:** Reviewers can use LLMs to assess clarity, logical structure, or even check for factual consistency in research papers.
6. **Hypothesis Generation:** LLMs can suggest novel hypotheses or experimental designs by identifying gaps in the literature.

A Key Thinker Behind the LLM: Jakob Uzkoreit

Uszkoreit co-authored the groundbreaking 2017 paper **“Attention Is All You Need,”** which introduced the Transformer architecture—a milestone that underlies almost all modern LLMs.

The leap that Uzkoreit made (along with his seven co-authors) was to abandon structured methods for language analysis and to predict just one word at a time in a sentence by associating context in that sentence (and adjoining text) with similar contexts in a large corpus of previously analyzed text.

At its core, the Transformer relies on a mechanism called self-attention, which allows each token in a sequence to weigh its relationships with all other tokens directly, regardless of distance using a **knowledge graph**. This makes it possible to capture long-range dependencies efficiently and in parallel. The paper introduced two key components: the encoder, which processes input sequences, and the decoder, which generates outputs (e.g., in machine translation). The architecture scales elegantly because the self-attention mechanism runs in $O(n2)$ time with respect to sequence length but can be highly parallelized on rapid graphic processing units from Nvidia.

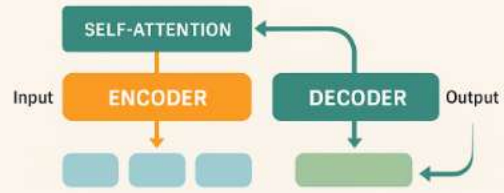
Source: https://www.wired.com/story/eight-google-employees-invented-modern-ai-transformers-paper/ and Stephen Witt, Thinking Machine: Jensen Huang, Nvidia, and the World's Most Coveted Microchip, 2025 (Chap 15).

Computational Linguistics and LLMs:

Why Products like ChatGPT, Google's Gemini, xAI's Grok and DeepSeek Are So Popular

Users that try out chat-based LLM products are typically **thrilled** with the results they get. It is as if the computer is thinking and generating results like a human and there is little constraint on the topic that one might ask about. Many users find results from popular LLM systems to be miraculous.

There are three key points to make as to why these systems are so good versus what was available before:



1. The transformer architecture engineered by Uzkoreit and others exploits the **context embedded in human language** to build a predictive model using the self-attention mechanism that transforms inputs via an encoder into output via a decoder. Language is laden with context and repetition which is leveraged.
2. The transformer architecture is relatively simple and massively **parallelizable** – which allows it to run parallel computing architecture such as that of Nvidia GPU farms in a manner that can process billions of web pages.
3. The implementations of this architecture from companies like OpenAI includes an **integrated routing system** that routes queries to the right processing point and applications including (a) reasoning engines, (b) multimodel input models) and (c) adaptive compute allocation which provides more computing power for complex queries.
4. The reasoning engine capability is particularly relevant for medical and scientific applications. DeepSeek and OpenAI's "thinking" model is built to explicitly perform internal chain-of-thought reasoning, breaking down problems step by step — an ability inspired by previous specialist reasoning models. (see https://cirra.ai/articles/gpt-5-technical-overview).

Claude deepseek Gemini OpenAI GPT-5 Grok perplexity

Today's LLM's Can Suffer From Human-Like Biases

LLMs are trained on human-produced data and thus inherit the same biases and limitations as humans.

Ethnicity and Gender Biases

Irene Chen and Emily Alsentzer, "Redefining Bias Audits for Generative AI in Health Care?," *NEJM AI*, Aug 14, 2025.

Large language models (LLMs) are transforming health care by supporting a range of administrative and clinical tasks; however, recent studies have raised concerns about their potential to exacerbate existing health care inequities. Traditional algorithmic auditing approaches fall short in addressing the unique challenges posed by LLMs, which process complex text-based inputs and generate human-like outputs. In this perspective, we examine current approaches for evaluating LLM bias in clinical settings, identifying key gaps in existing audit methodologies. We propose comprehensive guidelines for categorizing and detecting biases in LLM applications and illustrate their application through two real-world deployed systems — in-basket patient response drafting and mental health chatbots. Finally, we offer concrete recommendations for advancing LLM bias evaluation in a rapidly evolving technological landscape.

Cognitive Biases

Siduo Chen, "Cognitive Biases in Large Language Model based Decision Making: Insights and Mitigation Strategies," *Applied and Computational Engineering*, March 13, 2025.

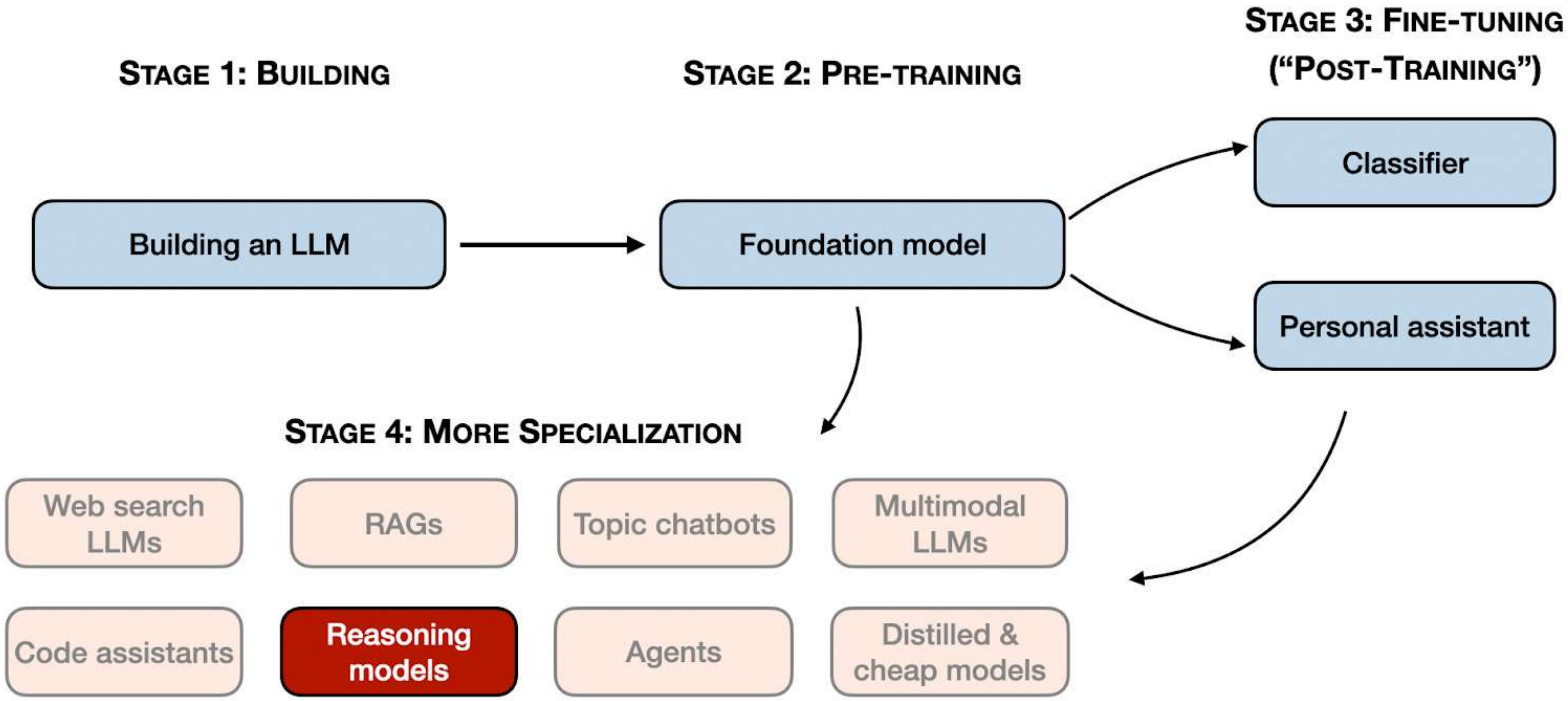
More recent research has shifted focus to the cognitive biases present in LLMs, a field that is relatively recent and less looked at. Initial works in this direction have successfully demonstrated underlying biases and heuristics including representativeness, insensitivity to sample size, base rate neglect, anchoring, and framing effects.

Alarmingly, LLMs themselves seem incapable or unaware of their fallacies, reassuring users of their impartialness when it comes to decision making. Successive research has since tested over 30 of the 180+ biases known to be liable in humans on LLMs and has paved the way for further research through providing a framework for defining and conducting tests, a dataset with 30,000 cognitive bias tests, and an in-depth evaluation of over 20 state-of-the-art LLMs.

Sources: https://ai.nejm.org/doi/abs/10.1056/AIp2500015 and https://www.ewadirect.com/proceedings/ace/article/view/21389.

LLM’s Becoming More Specialized in 2025 As We Build Out Better Reasoning Models

Sebastian Raschka, Understanding Reasoning LLMs, Blog Post, Feb 5, 2025



Source: https://magazine.sebastianraschka.com/p/understanding-reasoning-llms

It Is Possible to Build a Good Clinical Reasoning Model

Verification-Reinforced Reasoning Possible

Yun et.al., "Med-PRM: Medical Reasoning Models with Stepwise, Guideline-Verified Process Rewards," ***arXiv*****, June 13, 2025.**

Large language models have shown promise in clinical decision making, but current approaches struggle to localize and correct errors at specific steps of the reasoning process. This limitation is critical in medicine, where identifying and addressing reasoning errors is essential for accurate diagnosis and effective patient care. We introduce Med-PRM, a process reward modeling framework that leverages retrieval-augmented generation to verify each reasoning step against established medical knowledge bases. By verifying intermediate reasoning steps with evidence retrieved from clinical guidelines and literature, our model can precisely assess the reasoning quality in a fine-grained manner. Evaluations on five medical QA benchmarks and two open-ended diagnostic tasks demonstrate that Med-PRM achieves state-of-the-art performance, with improving the performance of base models by up to 13.50% using Med-PRM.

New Software Can Reason Well

Maziyar Panahi, "Introducing the First Commercially Available Medical Reasoning LLM," ***Blog Post*****, March 25, 2025.**

Our recent release of the Medical LLM suite demonstrated exceptional performance across critical benchmarks like MedQA, PubMedQA, and MedMCQA, with our 70B model achieving an impressive 86.2% average score across key medical knowledge assessments.

While these benchmarks confirm our models' exceptional medical knowledge, today's healthcare challenges demand more than just information retrieval—they require sophisticated reasoning capabilities that mirror clinical thinking. Medical practitioners face information overload, diagnostic uncertainty, and the constant risk of errors. Unlike models that focus primarily on producing direct answers, our reasoning models are designed to elaborate their thought processes—considering multiple hypotheses, evaluating evidence systematically, and explaining conclusions transparently.

Sources: https://arxiv.org/abs/2506.11474 and https://www.johnsnowlabs.com/introducing-the-first-commercially-available-medical-reasoning-llm/.

Three Types of Medical Reasoning Built into Software from John Snow Labs

John Snow LABS

Clinical reasoning is central to healthcare practice, encompassing the cognitive processes physicians use to evaluate patients, consider evidence, and make decisions. Our reasoning models are designed to emulate reasoning patterns in clinical practice.

Deductive Reasoning: Applying Rules to Specific Cases

- Deductive reasoning applies established principles or guidelines to individual cases – a foundational aspect of evidence-based medicine. For example, if a patient presents with fever, sore throat, and positive rapid strep test, clinical guidelines dictate oral penicillin as first-line treatment (for non-allergic patients). This reasoning follows a clear "if A and B, then C" logic flow that medical professionals use daily.
- When applied to healthcare AI, deductive reasoning enables models to systematically apply clinical guidelines, protocols, and established medical knowledge to specific patient scenarios. Medical Reasoning LLM can process a patient's symptoms, test results, and medical history, then apply relevant guidelines to recommend appropriate next steps – just as a clinician would when following clinical pathways.

Inductive Reasoning: Finding Patterns in Clinical Observations

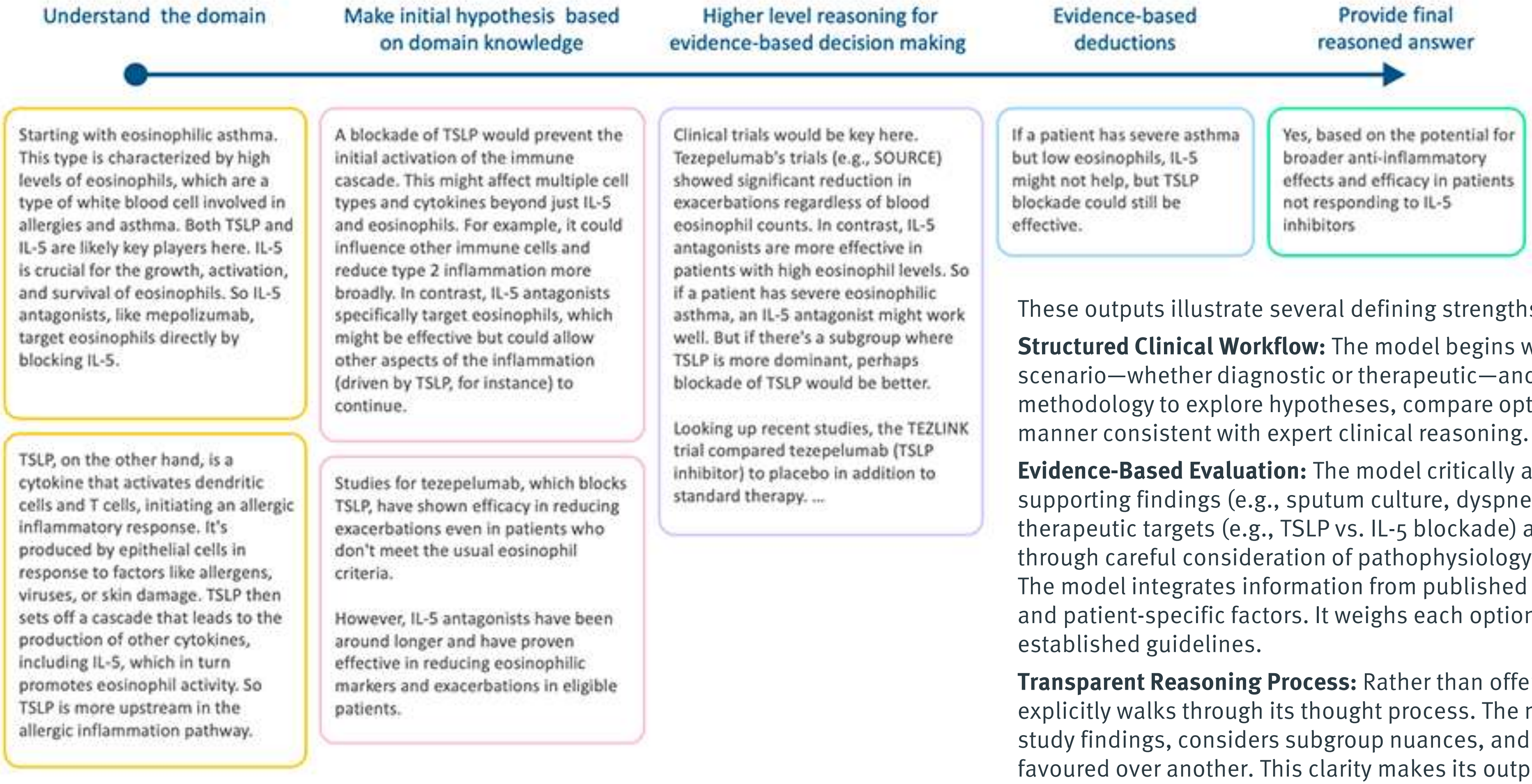
- Inductive reasoning moves from specific observations to general conclusions – vital for hypothesis formation in clinical settings. When an emergency physician notices that multiple unrelated patients present with similar respiratory symptoms within hours, they might inductively reason that an environmental factor like air quality or a viral outbreak could be affecting the community. This pattern recognition leads to broader hypotheses that inform both individual treatment and public health responses.
- In medical AI, inductive reasoning capabilities allow models to identify patterns across patient cases and generate hypotheses about underlying causes or connections. Medical Reasoning LLM can analyze clusters of symptoms or test results and suggest potential diagnoses based on pattern recognition, similar to how experienced clinicians develop clinical intuition through years of practice.

Abductive Reasoning: Making the Best Explanation with Incomplete Information

- Perhaps most critical in urgent care scenarios is abductive reasoning – making the most plausible inference with limited information. When a trauma patient arrives unconscious with unilaterally dilated pupils and declining neurological function, physicians must rapidly assess the most likely explanation (potentially intracranial hemorrhage) and take immediate action before confirmatory imaging is available.
- For medical AI, abductive reasoning is particularly valuable when dealing with incomplete patient information or time-sensitive decisions. Medical Reasoning LLM abductive reasoning capabilities allow it to suggest the most likely explanations for a given set of symptoms and prioritize critical actions, even when data is limited. The model can explicitly acknowledge uncertainty while still providing actionable insights based on the available information.

Source: https://www.johnsnowlabs.com/introducing-the-first-commercially-available-medical-reasoning-llm/.

IQVIA Reasoning Engine on an Illustrative Question: “Does blockade of TSLP offer greater clinical benefit in severe eosinophilic asthma compared to IL-5 antagonists?”



These outputs illustrate several defining strengths of the model’s reasoning capabilities:

Structured Clinical Workflow: The model begins with a clear articulation of the clinical scenario—whether diagnostic or therapeutic—and then applies a step-by-step methodology to explore hypotheses, compare options, and synthesize evidence in a manner consistent with expert clinical reasoning.

Evidence-Based Evaluation: The model critically assesses each diagnosis—identifying supporting findings (e.g., sputum culture, dyspnea, risk factors) or comparing therapeutic targets (e.g., TSLP vs. IL-5 blockade) and eliminating less likely options through careful consideration of pathophysiology, presentation, and contraindications. The model integrates information from published clinical trials, biological mechanisms, and patient-specific factors. It weighs each option using evidence-based reasoning and established guidelines.

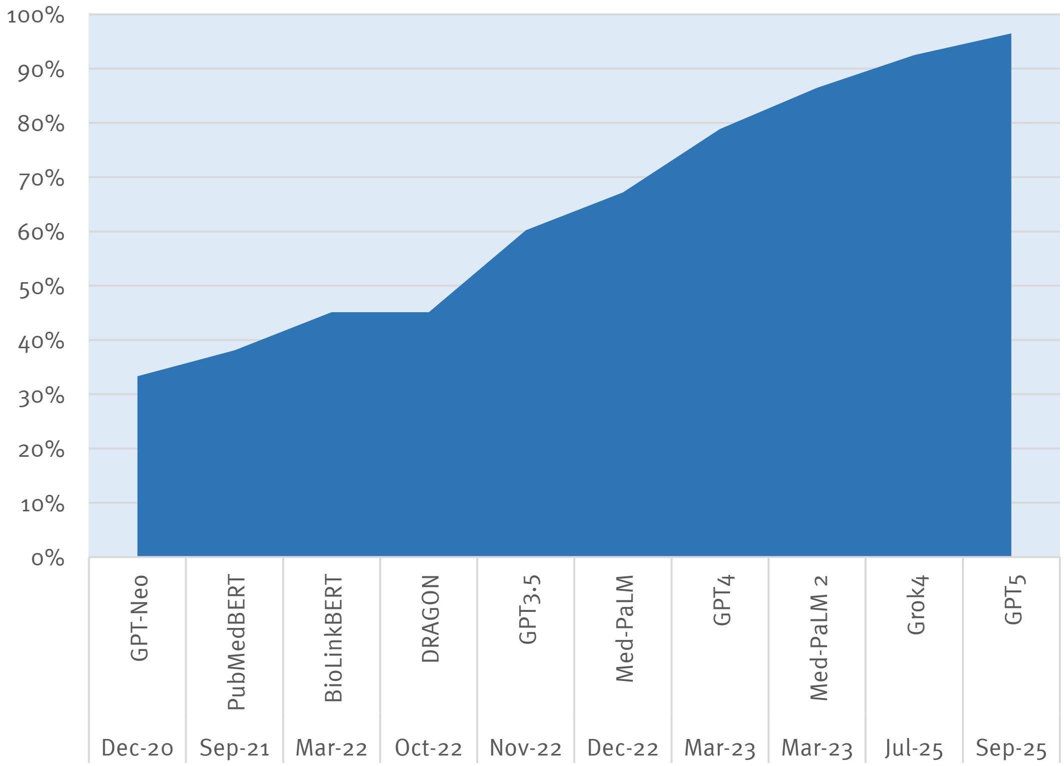
Transparent Reasoning Process: Rather than offering a conclusion outright, the model explicitly walks through its thought process. The model outlines assumptions, interprets study findings, considers subgroup nuances, and explains why one path may be favoured over another. This clarity makes its output traceable and verifiable.

Source: https://www.iqvia.com/blogs/2025/04/introducing-iqvia-medical-reasoning-med-r1-8b

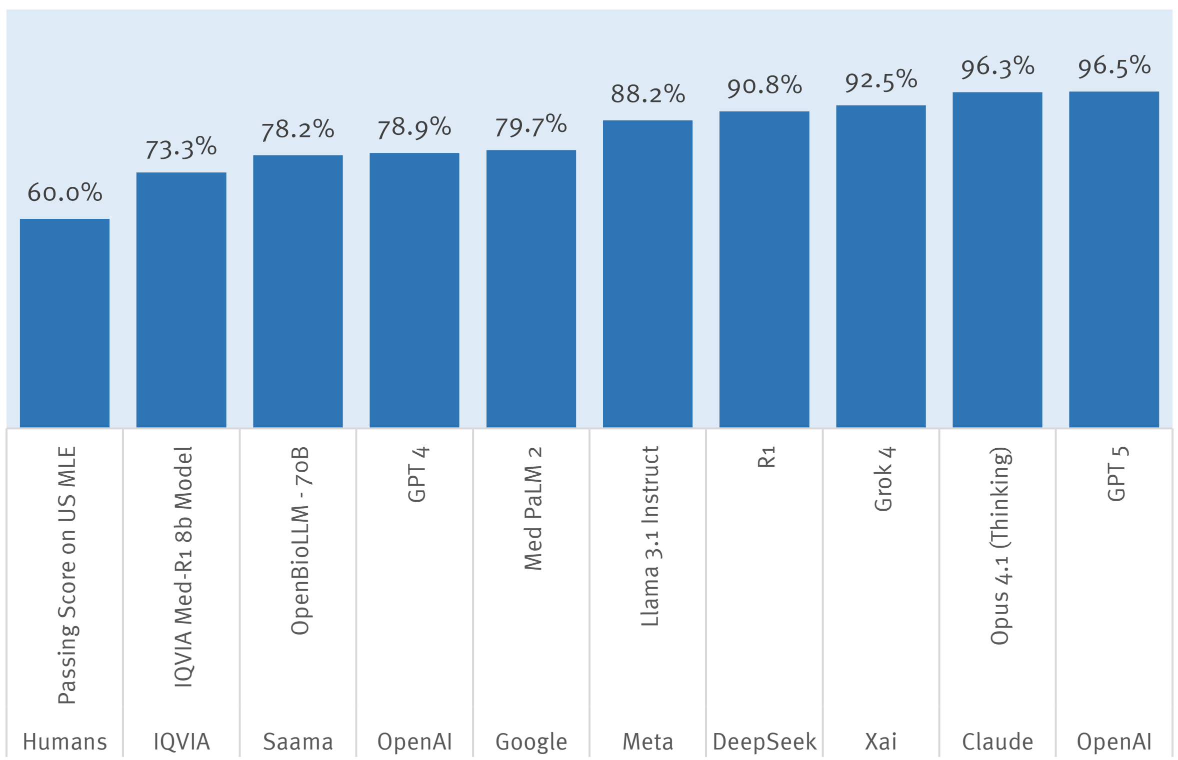
Benchmarking of AI Clinical Reasoning Systems

The MedQA dataset is a widely used biomedical question answering dataset that comprises questions in the style of the US Medical Licensing Exam (USMLE) and is used for evaluation of LLMs' reasoning capabilities. As is evident below, the systems have gotten progressively better to the point that today systems like Grok4 and GPT5 get almost no clinical questions wrong. These systems are far better than average doctors and beat expert doctors. While this may seem amazing, a human that could access all medical knowledge in less than a second with a strong reasoning system should also be able to attain a perfect score.

MedQA Benchmark with LLMs, 2020 to 2025
(Percent Questions Right on USMLE)



MedQA Benchmark with LLMs, Sep 2025
(Percent Questions Right on USMLE)



Source: Stifel research.

Surprisingly, AI Models Do Better When Physician Input is Not Included

By Pranav Rajpurkar and Eric J. Topol, Guest Editorial, *New York Times*, Feb 2, 2025 (excerpt)

The rapid rise in artificial intelligence has created intense discussions in many industries over what kind of role these tools can and should play — and health care has been no exception. The medical community largely anticipated that combining the abilities of doctors and A.I. would be the best of both worlds, leading to more accurate diagnoses and more efficient care.

That assumption might prove to be incorrect. A growing body of research suggests that A.I. is outperforming doctors, even when they use it as a tool.

A recent M.I.T.-Harvard study, of which one of us, Dr. Rajpurkar, is an author, examined how radiologists diagnose potential diseases from chest X-rays. The study found that when radiologists were shown A.I. predictions about the likelihood of disease, they often undervalued the A.I. input compared to their own judgment. The doctors stuck to their initial impressions even when the A.I. was correct, which led them to make less accurate diagnoses. Another trial yielded a similar result: When A.I. worked independently to diagnose patients, it achieved 92 percent accuracy, while physicians using A.I. assistance were only 76 percent accurate — barely better than the 74 percent they achieved without A.I.

This research is early and may evolve. But the findings more broadly indicate that right now, simply giving physicians A.I. tools and expecting automatic improvements doesn't work. Physicians aren't completely comfortable with A.I. and still doubt its utility, even if it could demonstrably improve patient care.

But A.I. will forge ahead, and the best thing for medicine to do is to find a role for it that doctors can trust. The solution, we believe, is a deliberate division of labor. Instead of forcing both human doctors and A.I. to review every case side by side and trying to turn A.I. into a kind of shadow physician, a more effective approach is to let A.I. operate independently on suitable tasks so that physicians can focus their expertise where it matters most.

Source: https://www.nytimes.com/2025/02/02/opinion/ai-doctors-medicine.html

In Study After Study Physician Input Degrades AI Decisions

Pranav Rajpurkar and Eric J. Topol, *Substack Blog*, Feb 2, 2025 (excerpt)

Surprisingly, in many cases, A.I. systems working independently performed better than when combined with physician input. This pattern emerged consistently across different medical tasks, from chest X-ray and mammography interpretation to clinical decision-making. In some of the studies, summarized in the Table, the gap for performance favoring A.I. alone was large.

A.I. Performance Exceeds Physician Plus A.I.?

Interpretative Task	AI Comparison	Study Design	Main Results AI vs MDs + AI	Citation
CXR	227 radiologists	CXR for 15 tasks	AI better than 2/3 radiologists w/AI	Agarwal, NBER, 2024
Mammography	14 radiologists	Cancer detection	0.94 vs 0.88	Kim HE, Lancet Digital Health, 2020
CXR	13 physicians	CXR for TB	0.79 vs 0.65	Rajpurkar P, NPJ Digital Medicine, 2020
Diagnostic accuracy	20 internists	RCT NEJM Cases	0.59 vs 0.52*	McDuff D, arXiv, 2023
Diagnostic accuracy	50 physicians	RCT GPT-4	0.90 vs 0.76	Goh E, JAMA Network Open, 2024
Management reasoning	25 physicians	GPT-o1, 5 cases	86% vs 41%^	Brodeur P, arXiv, 2024

*Top-10 accuracy; ^ median score o1 vs MDs plus GPT-4; RCT-randomized, controlled trial

CXR = Chest X-Ray

Source: https://erictopol.substack.com/p/when-doctors-with-ai-are-outperformed

Doctors Can Make Costly Diagnostic Errors

Charlotte Blease, "The Big Idea: why we should embrace AI doctors: People are understandably wary of new technology, but human error is often more lethal," *The Guardian*, August 31, 2025 (excerpt)

We expect our doctors to be demi-gods – flawless, tireless, always right. But they are only human. Increasingly, they are stretched thin, working long hours, under immense pressure, and often with limited resources. Of course, better conditions would help, including more staff and improved systems. But even in the best-funded clinics with the most committed professionals, standards can still fall short; doctors, like the rest of us, are working with stone age minds. Despite years of training, human brains are not optimally equipped for the pace, pressure, and complexity of modern healthcare.

Given that patient care is medicine's core purpose, the question is who, or what, is best placed to deliver it? AI may still spark suspicion, but research increasingly shows how it could help fix some of the most persistent problems and overlooked failures – from misdiagnosis and error to unequal access to care.

As patients, each of us will face at least one diagnostic error in our lifetimes. **In England, conservative estimates suggest that about 5% of primary care visits result in a failure to properly diagnose, putting millions of patients in danger. In the US, diagnostic errors cause death or permanent injury to almost 800,000 people annually.** Misdiagnosis is a greater risk if you're among the one in 10 people worldwide with a rare disease.

Modern medicine prides itself on being scientific, yet doctors don't always practise what the evidence recommends. Studies show that evidence-based treatments are offered only about half the time to adults in the US. Doctors can also disagree about diagnoses. In a study of more than 12,000 radiology images, reviewers offering second opinions disagreed with the original assessment in about one in three cases – leading to a change in treatment nearly 20% of the time. As the work day wears on, quality slips further: inappropriate antibiotic prescriptions rise, while cancer screening rates fall.

Source: https://www.theguardian.com/books/2025/aug/31/the-big-idea-why-we-should-embrace-ai-doctors

AI Models for Clinical Decision-Making Still Not Perfect

Buess, L., Keicher, M., Navab, N. et al., "From large language models to multimodal AI: a scoping review on the potential of generative AI in medicine," ***Biomedical Engineering Letters*** **15, September 2025, pp. 845–863.**

Multimodal LLMs extend LLM capabilities by integrating multiple data types, such as medical images and text, to address tasks like report generation, cross-modal retrieval, and clinical question answering. Despite these advancements, data heterogeneity remains a challenge, as clinical datasets often vary significantly in format, quality, and completeness across institutions.

Source: https://link.springer.com/article/10.1007/s13534-025-00497-1

Gourabathina, A. et.al., "The Medium is the Message: How Non-Clinical Information Shapes Clinical Decisions in LLMs," ***FAccT '25: Proceedings of the ACM*****, June 23, 2025**

Through the perturbation of patient messages, we evaluate whether LLM behavior remains consistent, accurate, and unbiased when non-clinical information is altered. These perturbations assess the brittleness of clinical LLM reasoning by replicating structural errors that may occur during electronic data processing patient questions and simulating interactions between patient-AI systems in diverse, vulnerable patient groups. Our findings reveal notable inconsistencies in LLM treatment recommendations and significant degradation of clinical accuracy in ways that reduce care allocation to patients.

Source: https://dl.acm.org/doi/10.1145/3715275.3732121

Imperfect or Not, there is Good Reason to Prefer AI over Your Local Doctor

Charlotte Blease, "The Big Idea: why we should embrace AI doctors: People are understandably wary of new technology, but human error is often more lethal," *The Guardian*, August 31, 2025 (excerpt)

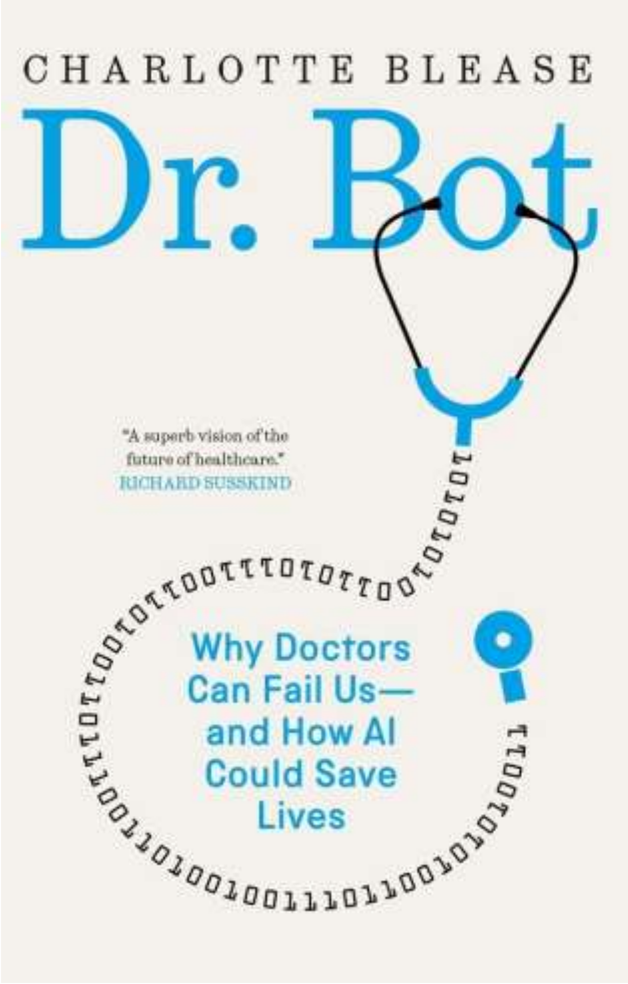
As alarming as this is, there are understandable reasons for these failures – and viewed from another angle, it's remarkable that doctors get it right as often as they do. The realities of being human – distraction, multitasking, even our body clocks – take a toll. But burnout, depression and cognitive ageing don't just wear doctors down; they raise the risk of clinical mistakes.

Medical knowledge also moves faster than doctors can keep up. By graduation, half of what medical students learn is already outdated. It takes an average of 17 years for research to reach clinical practice, and with a new biomedical article published every 39 seconds, even skimming the abstracts would take about 22 hours a day. There are more than 7,000 rare diseases, with 250 more identified each year.

In contrast, AI devours medical data at lightning speed, 24/7, with no sleep and no bathroom breaks. Where doctors vary in unwanted ways, AI is consistent. And while these tools make errors too, it would be churlish to deny how impressive the latest models are, with some studies showing they vastly outperform human doctors in clinical reasoning, including for complex medical conditions.

AI's superpower is spotting patterns humans miss, and these tools are surprisingly good at recognising rare diseases – often better than doctors. For example, in one 2023 study researchers fed 50 clinical cases – including 10 rare conditions – into ChatGPT-4. It was asked to provide diagnoses in the form of ranked suggestions. It solved all of the common cases by the second suggestion, and got 90% of the rare conditions by the eighth – outperforming the human doctors used as comparators.

Source: https://www.theguardian.com/books/2025/aug/31/the-big-idea-why-we-should-embrace-ai-doctors

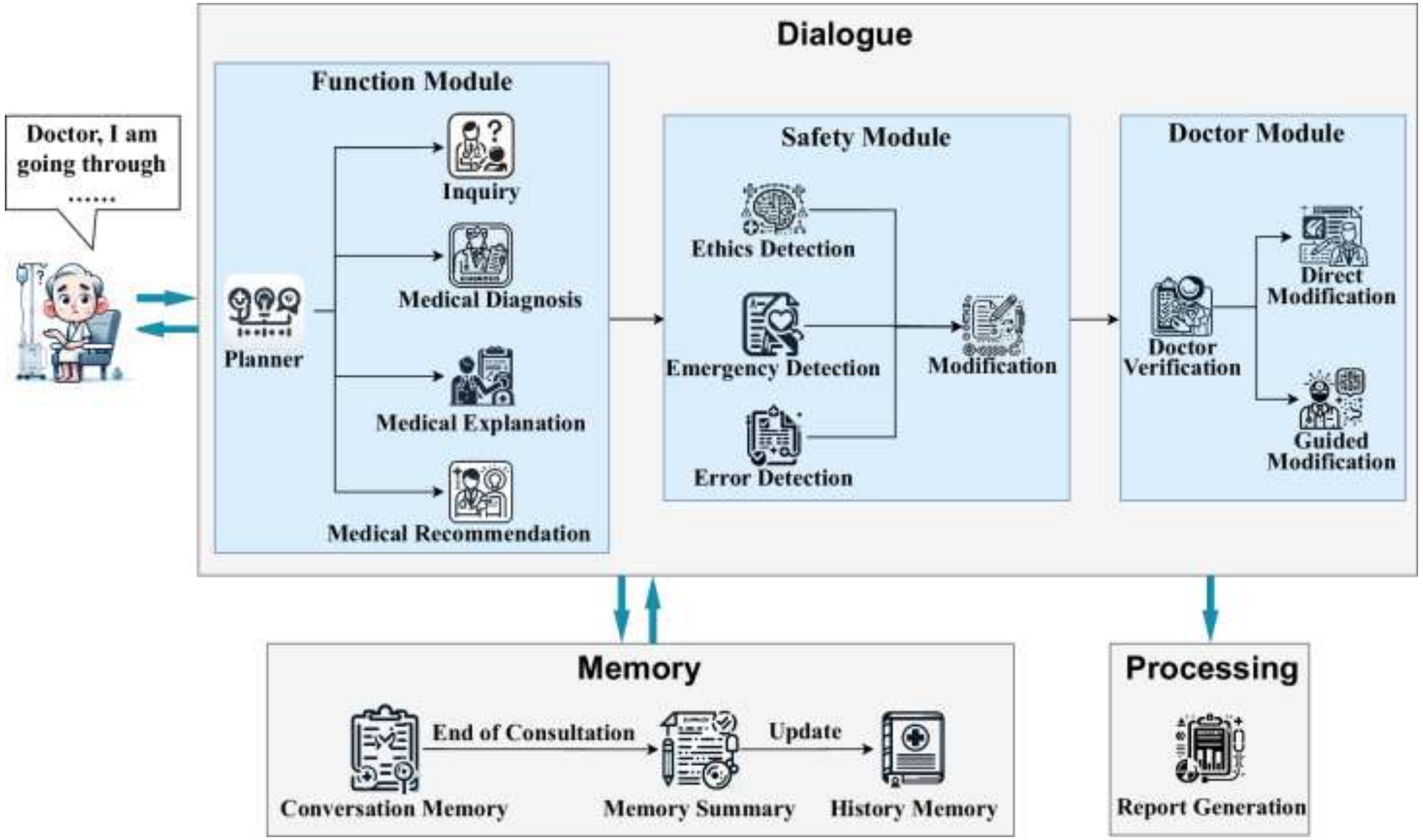


Charlotte Blease's book "Dr. Bot" was released on Sep 9, 2025.

Hybrid Models: Possible to Design a MultiModal Model that Combines LLM’s with Real World Physician Verification

Ren, Z., Zhan, Y., Yu, B. *et al.* Healthcare agent: eliciting the power of large language models for medical consultation. *npj Artif. Intell.* 1, 24, Sep 1, 2025.

Large language models (LLMs) have achieved remarkable success across many fields but face challenges in complex real-world scenarios like medical consultation, particularly regarding inquiry quality and safety concerns. In this paper, we introduce a healthcare agent designed to address these issues, which cannot be fully resolved through a vanilla one-time fine-tuning process. The healthcare agent includes three components: a dialogue component for planning safe and effective conversations, a memory component for storing patient conversations and medical history, and a processing component for report generation. To evaluate our healthcare agent’s performance in medical consultations, we employ both expert assessment from medical professionals and an automated evaluation system powered by ChatGPT for large-scale testing. Our results demonstrate that the healthcare agent significantly enhances the capabilities of general LLMs in medical consultations, particularly in inquiry quality, response quality, and safety.



Source: https://www.nature.com/articles/s44387-025-00021-x

Microsoft's Competing Agent Approach Holds Promise

Melissa Heikkilä and Stephen Morris, *Financial Times*, June 30, 2025 (excerpt)

Microsoft has built an artificial intelligence-powered medical tool it claims is four times more successful than human doctors at diagnosing complex ailments, as the tech giant unveils research it believes could speed up treatment.

The "Microsoft AI Diagnostic Orchestrator" is the first initiative to come out of an AI health unit formed last year by Mustafa Suleyman with staff poached from DeepMind, the research lab he co-founded and which is now owned by rival Google.

In an interview with the *Financial Times*, the chief executive of Microsoft AI said the trial was a step on the path to "medical superintelligence" that could help solve staffing crises and long waiting times for overstretched health systems.

Microsoft's new system is underpinned by a so-called "orchestrator" that creates virtual panels of five AI agents acting as "doctors" — each with a distinct role, such as coming up with hypotheses or choosing diagnostic tests — which interact and "debate" together to choose a course of action.

To test its capabilities, "MAI-DxO" was fed 304 case studies from the *New England Journal of Medicine* (NEJM) that describe how some of the most complicated cases were solved by doctors. This allowed researchers to test if the programme could figure out the correct diagnosis and relay its decision-making process, using a new technique called "chain of debate", which makes AI reasoning models give a step-by-step account of how they solve problems. The orchestrator made all LLMs perform better, but worked best with OpenAI's o3 reasoning model to correctly solve 85.5 per cent of the NEJM cases.

That compared with about 20 per cent by experienced human doctors, but those physicians were not allowed access to textbooks or to ask colleagues in the trial, which could have increased their success rate.

Source: https://www.ft.com/content/149296b9-41b6-4fba-b72c-c72502d01800

Physicians are Rapidly Adopting LLM Tools at Work

Elsevier Health Report, *Clinician of the Future*, June 2025.

Of the 48% of clinicians who use AI for work, nearly all (97%) have used a generalist AI tool, such as OpenAI's ChatGPT or Google's Gemini (formerly Bard). Usage of these tools has increased in the past year: in 2024, only 8% of clinicians used AI tools frequently (saying they used AI a lot); in the current survey, 50% use these tools frequently (41%) or always (9%).

In comparison, 76% of those who have used AI for work have used a clinical-specific AI tool. Only 18% use them frequently and 4% always.

March / April 2025 survey of 2,206 clinicians from 109 countries

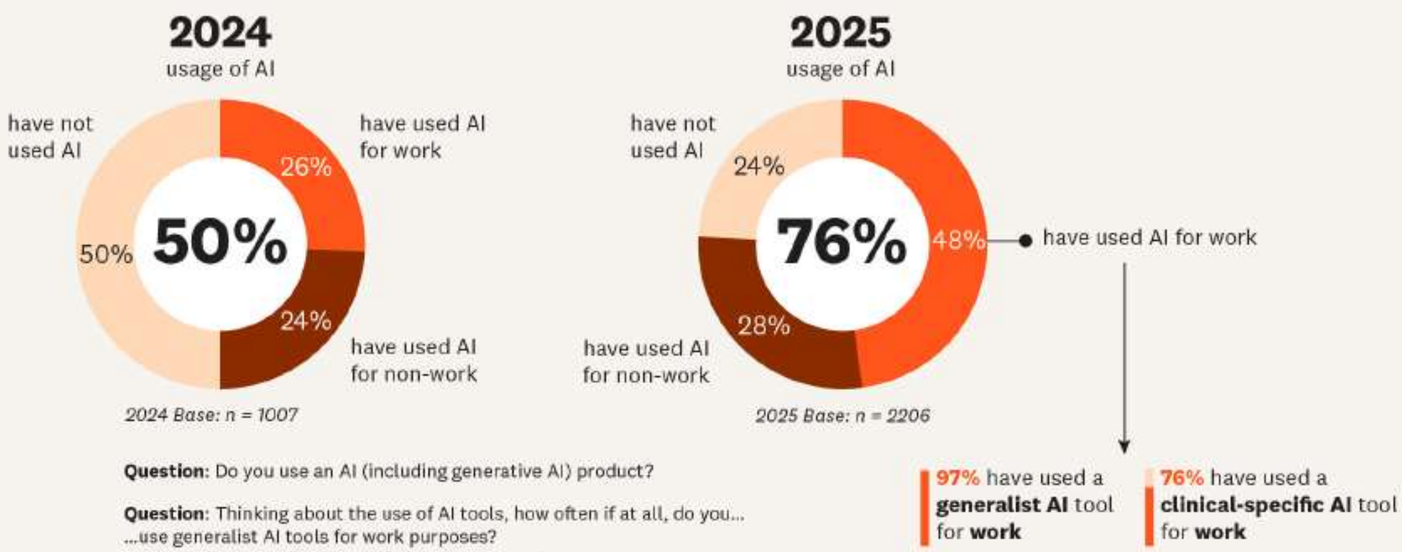
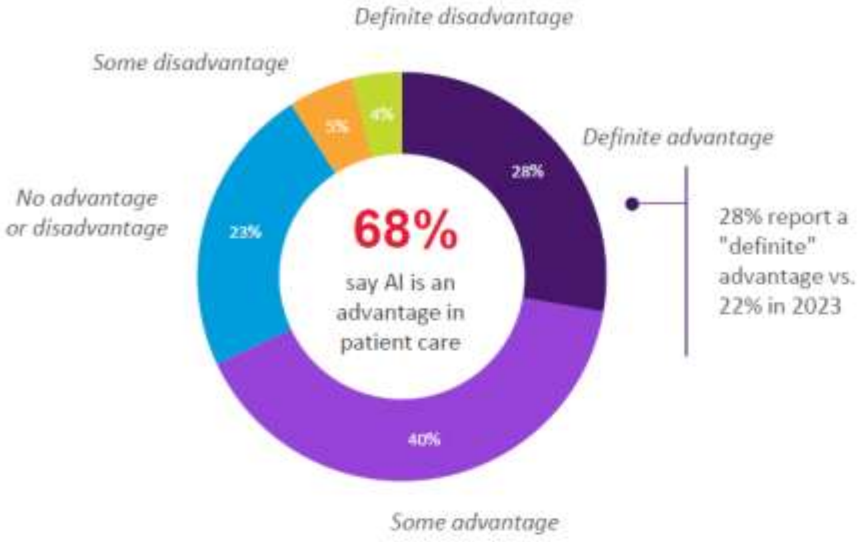


Figure 4

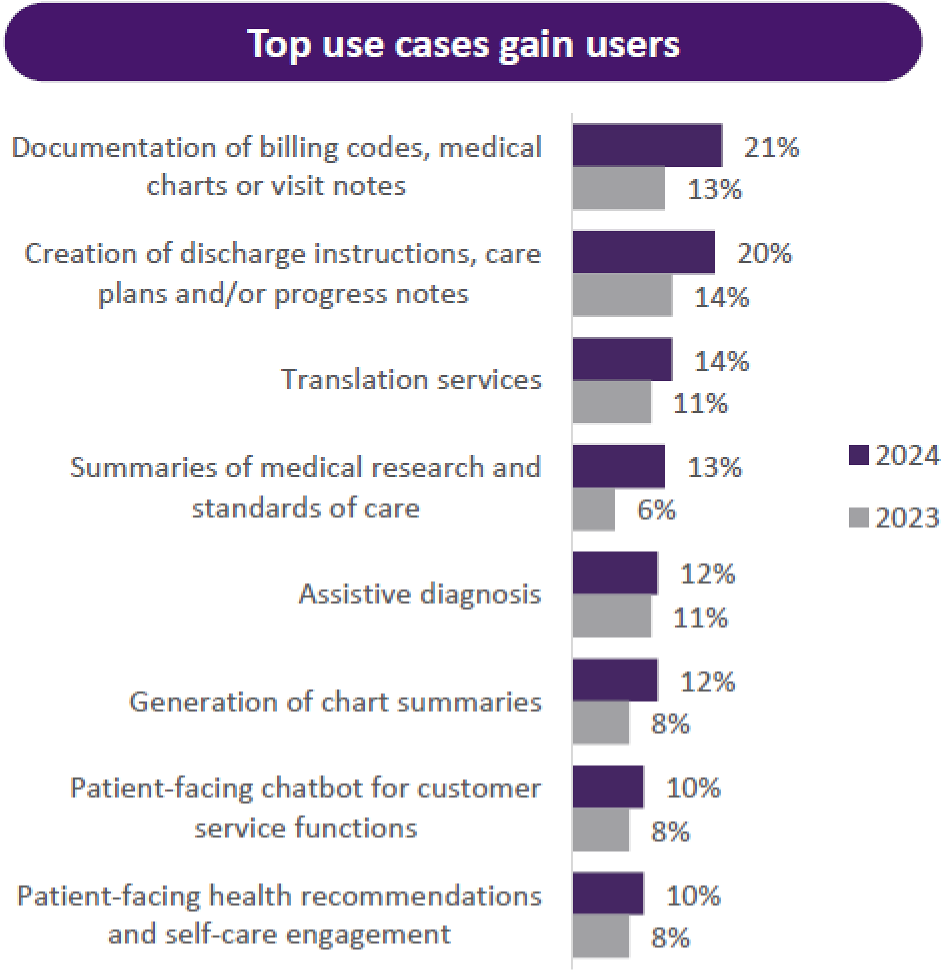
Source: https://www.elsevier.com/insights/clinician-of-the-future/2025

Physician Confidence and Use of AI is Rising Fast

AMA Augmented Intelligence Research: Physician sentiments around the use of AI in heath care: motivations, opportunities, risks, and use cases - Shifts from 2023 to 2024, February 2025



Percent of respondents stating they currently use at least one of the 15 AI use cases presented



Source: https://www.ama-assn.org/system/files/physician-ai-sentiment-report.pdf

Decision Support Tools That Combine LLM and Reasoning Engines for Physician Use

Various specialized tools are being variously designed to integrate imaging and diagnostics, to reduce administrative burden, to be HIPAA compliant, to integrate with existing EHR's like EPIC and to build models only from trusted sources like the NEJM. The most widely used tool appears to be OpenEvidence which provides better sourcing that ChatGPT and just raised $210 million at a $3.5B valuation led by Sequoia & Kleiner Perkins.

Developer	Model / Tool	Main Function	Key Features / Performance
Google Research	Med-PaLM 2 / Med-PaLM M	Medical QA & reasoning (text + multimodal)	Med-PaLM 2 scores ~86.5% on MedQA; preferred over physicians in some long-form answers; Med-PaLM M adds image (X-ray, pathology) support.
M42	Med42-v2	Open, clinically aligned LLMs	Outperforms base Llama3 and GPT-4 on medical benchmarks; optimized for real clinical prompts.
HIPPOCRATES	Hippocrates / Hippo Models (open source)	Framework + LLM family for medical use	~7B-parameter models with instruction tuning & RLHF; transparent and democratized medical AI.
AVO	Ask Avo	Clinical decision support for physicians	Retrieval-augmented with citations; scored higher than GPT-4 in trustworthiness, actionability, and relevance.
OpenEvidence	OpenEvidence & DeepConsult	AI medical literature search & synthesis	Used by ~40% of U.S. physicians daily; cites across 35M+ publications; DeepConsult generates research briefs for complex questions.
Google Research	Med-Gemini	AI scribes for documentation	Generate clinical notes from physician–patient conversations; reduce clerical burden; HIPAA compliant.
ada	Assess by Ada	Symptom checker that physicians can use	A study evaluated Ada DX in rare disease diagnosis. Ada's top suggestion matched the confirmed diagnosis in 89% of cases (83 of 93 cases).
heidi	Heidi	AI scribe integrated with EHRs	Millions of weekly interactions; integrates with Epic, Athena; HIPAA compliant.
doctronic	Doctronic	AI diagnostic assistant for personal use + physician follow-up	Provides 4 differential diagnoses + summaries; 70% accuracy vs final physician Dx; 10M+ consultations; HIPAA-compliant.
Microsoft Research	MAI-DxO (AI Diagnostic Orchestrator)	Multi-agent diagnostic reasoning	"Chain-of-debate" AI agents; 85.5% success on NEJM complex cases vs ~20% for physicians; still research-stage.

Hundreds of Specialty AI Applications in Development

We are seeing a proliferation of specialty medical artificial intelligence applications. Key themes include earlier detection, workload reduction, making the right decisions and preventative health. Here are some examples of current applications in use.

Tool / Company	Domain	Key Use Case	Notable Features / Impact
AiCure	Digital therapeutics / Adherence monitoring	Smartphone-based AI visual recognition ensures patients take medications correctly	Used in clinical trials; improves adherence and real-world evidence quality
Ambience	Clinical documentation	AI scribe that generates notes from physician–patient encounters	HIPAA compliant; reduces clerical workload and burnout
LUMINETICSCORE AI Diagnostic System Formerly IDx-DR	Ophthalmology / Screening	FDA-approved autonomous AI for diabetic retinopathy detection	Deployed in primary care; allows non-specialists to screen for sight-threatening disease
Cambridge University Health Partners	Oncology (radiotherapy)	Automates radiotherapy treatment planning	Cuts planning from hours to minutes; improves efficiency in cancer care
PathAI	Pathology / Diagnostics	AI for pathology slide review and disease classification	Improves accuracy in cancer and liver disease diagnosis; used in research and pharma
TEMPUS	Precision oncology	AI + genomics platform for personalized cancer treatment	Combines molecular profiling with clinical AI to guide therapy selection
Viz.ai	Emergency neurology / Cardiology	AI triage and workflow software for stroke, aneurysm, cardiac emergencies	FDA-cleared; speeds specialist notification, improves time-to-treatment
Imperial College London	AI Stethoscope for Cardiology / Primary care	Detects AFib, valve disease, heart failure within ~15 seconds	Enables earlier detection in primary care and community settings

Rapid Consumer Adoption of LLM Guided Self-Care and Self-Diagnosis is Underway

Percent of US consumers who had diagnosed themselves with a commercially available LLM (2023 – 2025)

Early 2023	2023 to 2024	January 2025	August 2025
0%	21%	70%	> 80%

While healthcare systems debate whether to implement AI tools for clinical decision-making the consumer in many countries, particularly the U.S., has increasingly taken matters into their own hands. The consumer has gone from being suspicious of LLMs to having similar trust in them as physicians themselves. Our industry conversations indicate that primary care service utilization is dropping rather quickly in some of the more tech-savvy cities in the U.S. as a result. Our conversations with consumers are indicating high satisfaction because they are able to obtain relevant health information regarding their complaints and illnesses on demand and, so, are becoming more engaged with their own health matters.

Sources: (1) 2023 to 2024: https://www.jmir.org/2025/1/e68560, https://www.bain.com/insights/call-the-doctor-are-patients-ready-for-generative-ai-in-healthcare-snap-chart/; (2) January 2025: https://www.bain.com/insights/pharma-commercialization-in-the-age-of-ai-and-active-patients/; (3) September 2025, Stifel estimate based on informal survey of a broad set of American people (N=40).

How AI Can Transform the Delivery of Medicine

There are many obvious and interesting applications of AI in medicine. We see three ways to go in the future: (1) doctors hold on to their current role and use AI themselves, (2) patients shift to self-care for routine needs and only see doctors for more severe specialty situations (e.g., heart failure or cancer) or (3) patients shift to a hybrid model where the computer is involved more continually in care but doctors oversee the bots to make sure that the right decisions are getting made.

Self Care (No Doctors)

Patients go from “Dr. Google” to a much better curated AI solution that escalates for them when they need it. Many patients have been using LLM’s in 2025 and have been quite impressed.

Patients with Poor Access

Sicker patients who aren’t going to be able to manage interaction with an app can consult with nurses by phone or technicians who pay home visits to sort out what to do, collecting key diagnostic data before making choices.

Auto Care (Hybrid Model)

Patients with routine needs respond to questions. The computer analyzes the situation, recommends prescriptions while the doctor reviews and approves them (or asks to see the patient). This is working quite well now in some settings.

Doctors Stay in Charge and Use AI Themselves

Today, physicians hit online sources, colleagues and the library to sort out what to do in difficult cases. An AI engine can be incredibly helpful in this case if properly prepared. A similar approach can be used in specialty care for advanced disease like cancer.

Our View: Hybrid Model Will Evolve From the Status Quo

The use of AI in medicine will not put physicians out of work at all but will increase the premium on knowledge of medicine and diagnostic technologies. We see a world with fewer physicians who oversee AI system and healthcare workers.

Master Physicians

1. Physicians with both deep clinical experience and/or strong understanding of bioinformatics, 'Omics and biological research. Substantial diversity here.
2. These physicians help to review data and appropriateness of algorithmic approaches to disease but do not necessarily call all the shots.
3. They are available for complex cases that get escalated.

Clinicians

1. A clinician can be either a nurse or a physician.
2. The key is experience in positively interacting with patients and understanding disease state etiology and their treatment.
3. These persons will have real-time access to algorithmic /AI type advice based on patient interactions – typically using something like an iPad.
4. A good clinician will be trained to escalate tougher cases rapidly.

Medical Technicians

1. Medical technicians are already becoming the first line of encounter with patients in many ambulatory settings like urgent care clinics. This will continue.
2. A well-trained medtech armed with an LLM-linked iPad in the future should be able to make much more progress than can be done today with Epic, for example.
3. A good medtech will rapidly escalate complex cases to a clinician.

Today's LLM is the First Step on a Long Future Path of Artificial Intelligence Applied to Medicine

Computers can analyze existing text to associate "good answers" to medical questions.

This capability is already available.

Text based LLM's are able to gain access to patient level data and imagery to customize answers.

This capability is in the works.

Computers gain increased ability to reason through clinical problems that have not been described before.

This capability is well underway.

Computers analyze deeper datasets including proteomics and genomics to give customized advice.

This capability is still in the future.

Where Medicine is Going From Here: Our View

Improved Understanding of Biology

What matters is how biologic systems work together (or not) to lead to death and aging. Much less about organ systems and more about cells, genes and proteins.

Today's view of "disease" and "health" will change to "where's your biology"?

Improved Ability to Measure Biology with a Blood Sample

With a much deeper understanding of core organismal molecular biology, it will also be possible to learn what is going with a patient leveraging new techniques

We can now use proteomics to measure a person's biological status

Improved Ability to Customize Interventions

Statistical analysis of measures of underlying biology with patient outcomes and the role of customized interventions will improve massively over time.

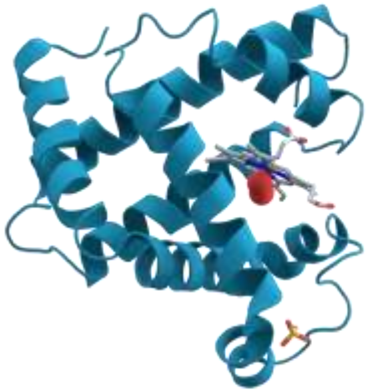
The role of AI is going to be critical. Machine learning can be quite helpful here.

Ability to Deliver Solutions to Improve Health

We are already seeing AI engines clobber doctors in clinical decision-making. Google's engine and DeepSeek both look surprisingly good.

We think of a physician of the future as a "pilot flying with instruments" – we are not getting rid of the airplane pilot. Yet.

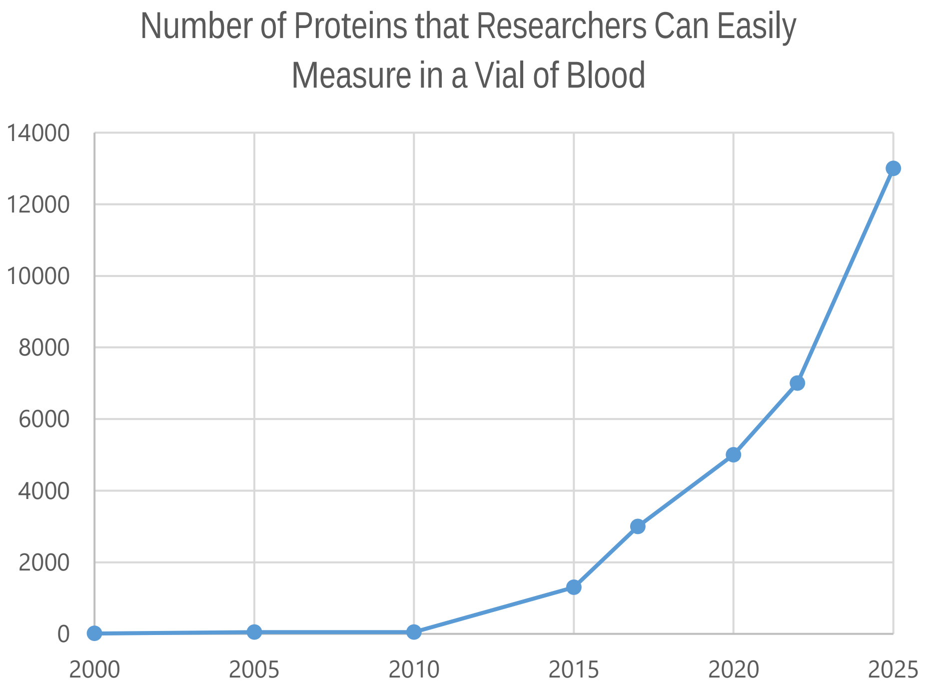
Proteins



express biology

Modern proteomic analysis allows us to gain a far deeper understanding of a person's health than was previously possible. While genetics are fixed at birth, proteins can be far more informative over our lifespan because they reflect environmental and physical conditions.*

New technologies allow scientists to measure thousands of proteins from a vial of blood.



Key Technologies:

olink

seer

somalogic

* See PLoS One. 2016 Apr 22;11(4):e0154387; *Aging Cell*. 2010 Dec;9(6):1057-64. *Commun Biol*. 2021 Jun 18;4(1):758.

Modern Proteomics: A Simple Blood Draw Can Give a Very Good Read of Health

- In a June 2021 research paper scientists from deCODE, a subsidiary of Amgen, describe a predictor of human mortality.
- Using a dataset of ~5000 protein measurements in 22,913 Icelanders, the scientists developed a predictor of life expectancy that substantially outperformed predictors based on traditional risk factors.
- "The predictor gives a good estimate of general health from a single blood draw," says Thjodbjorg Eiriksdottir scientist at deCODE genetics.

deCODE genetics
a subsidiary of AMGEN

Kari Stefansson and Thjodbjorg Eiriksdottir, deCODE

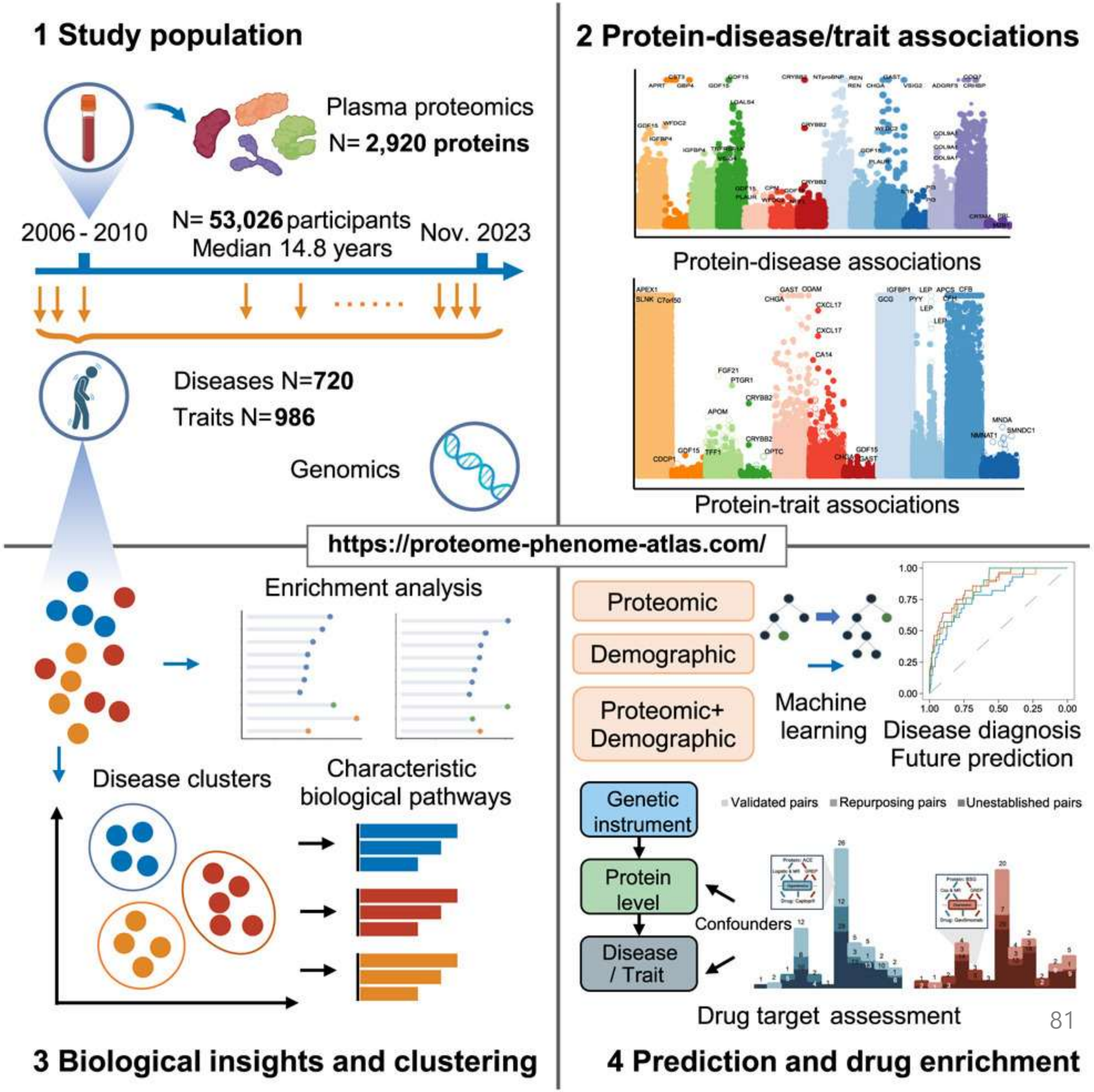
See: https://www.decode.com/predicting-the-probability-of-death/

Fudan University Team Uncovers Numerous Disease Biomarkers Using Mass Proteomics Machine Learning Exercise

Deng et.al., "Atlas of the plasma proteome in health and disease in 53,026 adults," *Cell*, Jan 9, 2025 pp. 253-271.e7.

Large-scale proteomics studies can refine our understanding of health and disease and enable precision medicine. Here, we provide a detailed atlas of 2,920 plasma proteins linking to diseases (406 prevalent and 660 incident) and 986 health-related traits in 53,026 individuals (median follow-up: 14.8 years) from the UK Biobank, representing the most comprehensive proteome profiles to date. This atlas revealed 168,100 protein-disease associations and 554,488 protein-trait associations. Over 650 proteins were shared among at least 50 diseases, and over 1,000 showed sex and age heterogeneity. Furthermore, proteins demonstrated promising potential in disease discrimination (area under the curve [AUC] > 0.80 in 183 diseases). Finally, integrating protein quantitative trait locus data determined 474 causal proteins, providing 37 drug-repurposing opportunities and 26 promising targets with favorable safety profiles

Source: https://www.cell.com/cell/fulltext/S0092-8674(24)01268-6



Mass Proteomics Technologies Enables Better Diagnostics Via AI

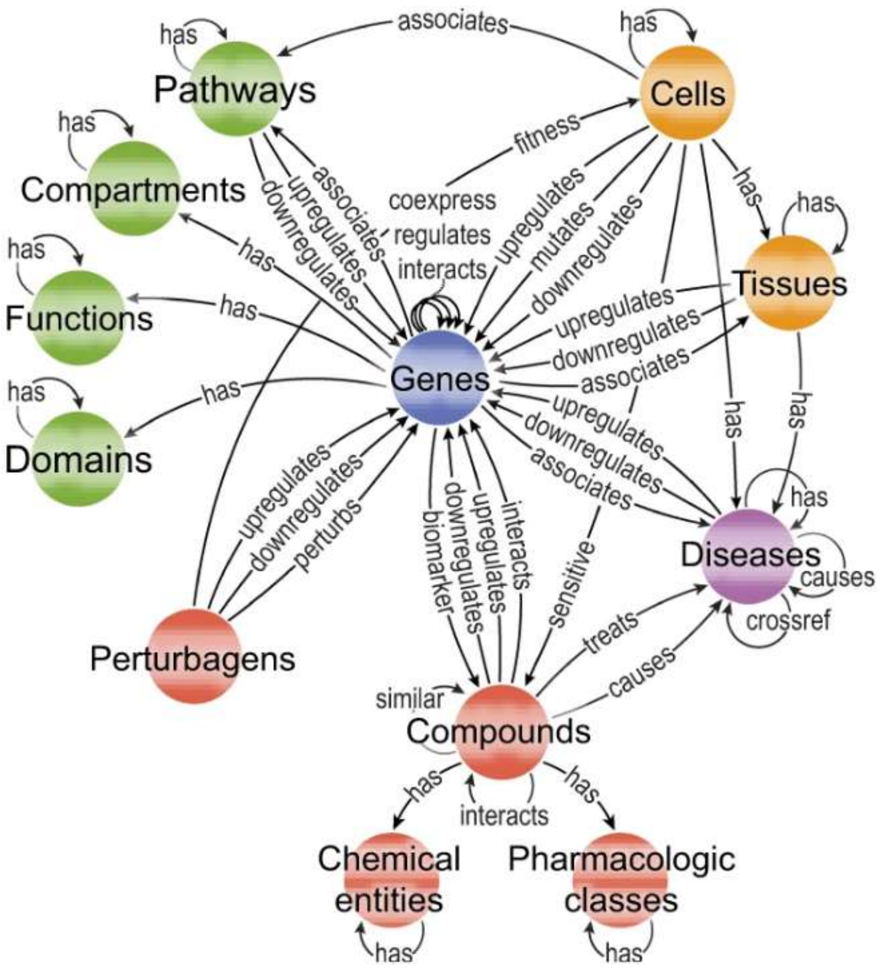
- It's possible to build much better disease diagnosis algorithms than those which are available today.
- Most of the factors that associate with diseases in the Decode studies involve subclinical pathologies, including ECM and cellular stress. None of the factors are in current FDA-approved tests.
- If one could diagnose disease more accurately and more quickly, there are major benefits for patients and payors.
- In the same sense that GRAIL built a pan-cancer detection algorithm it may be possible to build an inexpensive pan-disease detection algorithm using modern proteomics and metabolomics.
- We believe that the input from such analyses would be well suited for AI-based clinical decision-making tools. Already, we are seeing major progress in this area from companies like Health Outlook.
- A key issue is that modern methods could, in theory, overwhelm a physician with reams of data – that would not be challenging for an AI system to analyze against patterns seen in previous patients.

Medical Data is Made for Generative AI Applications

From a mathematical perspective, using neural networks to analyze clinical inputs and predict a patient's current and future status works in much the same way as a large language model (LLM).

Just as LLMs derive their power from the richness of human language and its consistency of use within cultural groups, human biology also shows broad consistency across individuals, with highly contextualized and interconnected systems.

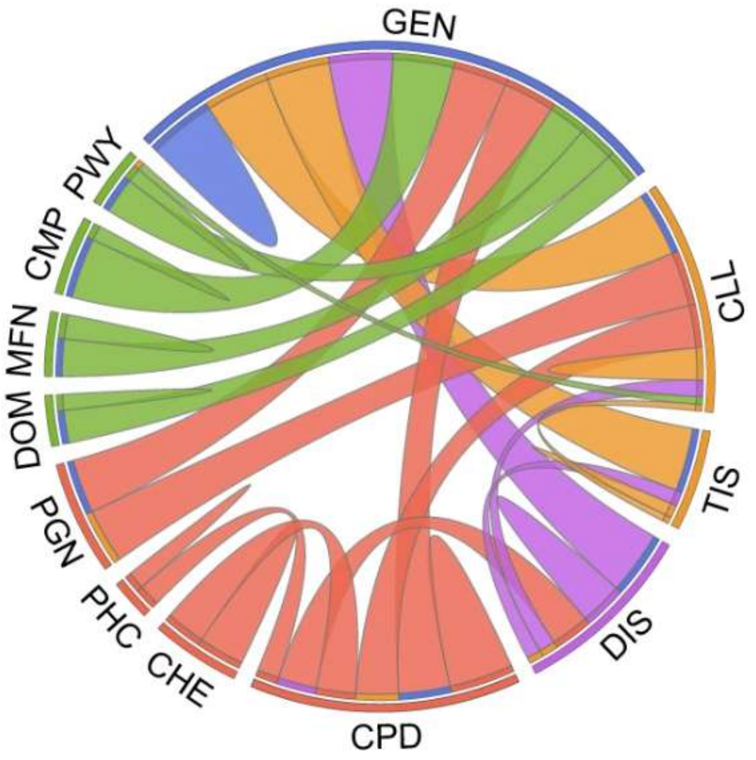
There are, however, important distinctions. Many health variables are continuous rather than discrete like words and digitizing them can generate extremely large datasets. In addition, biological relationships are strongly influenced by gender, ethnicity, and age—dependencies that do not arise in the same way in the language domain.



Possible to Glean a Vast Amount of Clinically Relevant Input From a Clinical Knowledge Graph That Integrates Proteomic Data with Medical Literature and Traditional Diagnostic Markers

Santos, A., Colaço, A.R., Nielsen, A.B. *et al.*, "A knowledge graph to interpret clinical proteomics data," *Nature Biotechnology* 40, 692–702 (2022).

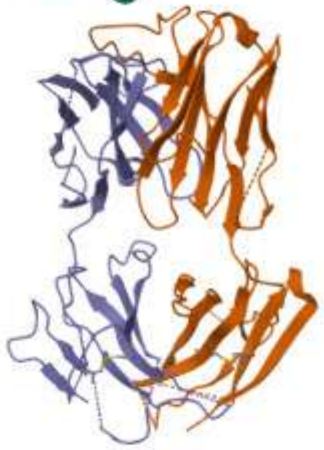
Implementing precision medicine hinges on the integration of omics data, such as proteomics, into the clinical decision-making process, but the quantity and diversity of biomedical data, and the spread of clinically relevant knowledge across multiple biomedical databases and publications, pose a challenge to data integration. Here we present the Clinical Knowledge Graph (CKG), an open-source platform currently comprising close to 20 million nodes and 220 million relationships that represent relevant experimental data, public databases and literature. The graph structure provides a flexible data model that is easily extendable to new nodes and relationships as new databases become available. The CKG incorporates statistical and machine learning algorithms that accelerate the analysis and interpretation of typical proteomics workflows. Using a set of proof-of-concept biomarker studies, we show how the CKG might augment and enrich proteomics data and help inform clinical decision-making.



Source: https://www.nature.com/articles/s41587-021-01145-6

Seeking Disease Insights, UK Biobank Launches Largest Study Yet of Human Proteins

Catherine Offord, *Science*, Jan 9, 2025.



Scientists are beginning the world's largest study of proteins circulating in human blood in an effort to better understand the development and treatment of disease.

Launched today as part of an ongoing collaboration between the British health study UK Biobank and more than a dozen pharmaceutical companies, the project aims to measure levels of 5400 proteins in blood samples from half a million people. For some of those people, researchers will analyze two separate samples taken years apart, generating a "first-of-its-kind" database on how a person's changing protein levels influence disease later in life, organizers say. The companies backing the project with tens of millions of dollars and carrying out the initial analysis of **300,000 blood samples** will get an exclusive 9-month window of access to the data, after which the information will be made more widely available to UK Biobank–approved research teams.

Scientists not involved in the project say the vast amount of data generated could aid development of blood tests that detect disease before symptoms appear and identify new drug targets for illnesses. "This will be an exceptionally powerful resource for understanding health and disease," says Eleftheria Zeggini, director of the Institute of Translational Genomics at Helmholtz Munich.

Source: https://www.science.org/content/article/seeking-disease-insights-uk-biobank-launches-largest-study-human-proteins

Low-Cost Large Scale Proteomics Platforms on the Horizon: Mass Spec Approaches

Huang et.al., "The $10 proteome: low-cost, deep and quantitative proteome profiling of limited sample amounts using the Orbitrap Astral and timsTOF Ultra 2 mass spectrometers," *bioRxiv*, July 31, 2025.

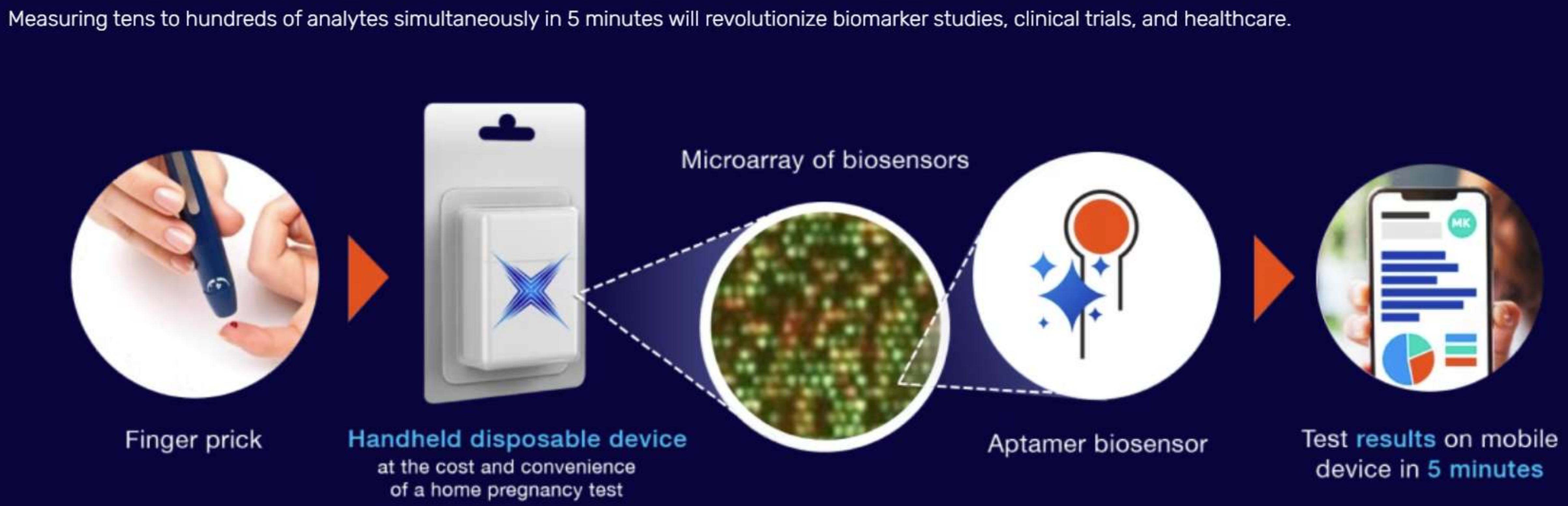
Mass spectrometry (MS)-based proteomics remains technically demanding and prohibitively expensive for many large-scale or routine applications, with per-sample costs of hundreds of dollars or more. To democratize access to proteomics and facilitate its integration into more high-throughput multi-omic studies, we present a robust analytical framework for achieving in-depth, quantitative proteome profiling at a cost of approximately $10 per sample, termed the "$10 proteome." Using the Thermo Fisher Orbitrap Astral and Bruker timsTOF Ultra 2 mass spectrometers, we evaluated performance across sample inputs ranging from 200 pg to 100 ng and active gradient lengths from 5 to 60 minutes. Proteome coverage saturated within the low-nanogram input range, with ~8000 protein groups quantified from as little as 10 ng of input and nearly 6000 protein groups from 200 pg.

The Thermo Fisher Orbitrap Astral Mass Spectrometer May Help Enable the $10 Proteome

Source: https://www.biorxiv.org/content/10.1101/2025.07.29.667408v1.full

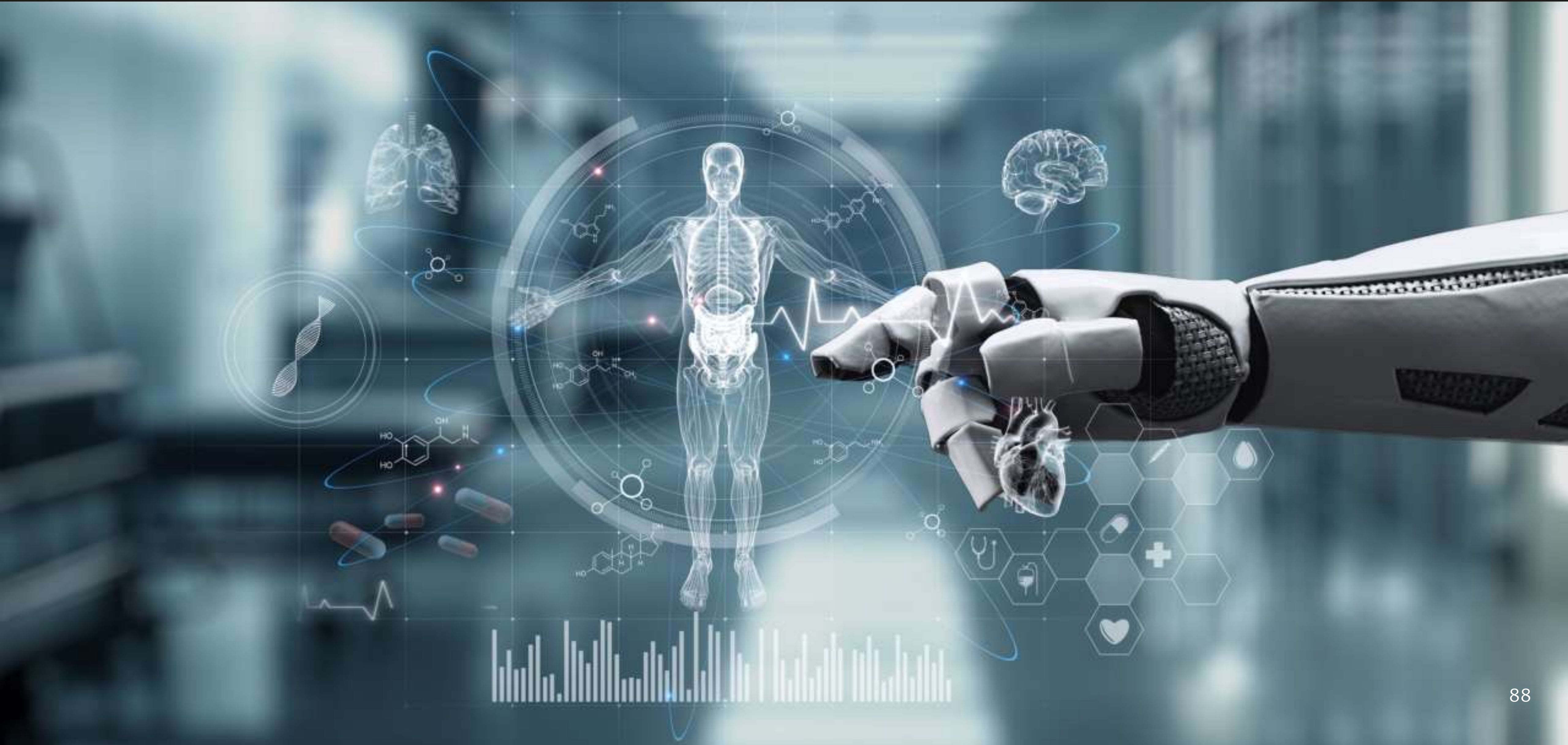
Low-Cost Large Scale Proteomics Platforms on the Horizon: eGLINT Proteonex Technology

eGLINT has implemented a technology that can measure major parts of the proteome for $5 or less. The benefit of this approach is that involves far less capital cost than the mass spec approach.



Source: https://www.eglint.com/

Might AI Speed Adoption of Innovation?



Humanity Not Doing a Great Job in Helping Humanity

To summarize what we have said so far is that human beings often find solutions to serious medical problems that are very slow to be adopted.

We have attributed this to lack of knowledge, social opposition, financial incentives, inability to communicate and ignorance.

The literature on innovation adoption basically says that you have to communicate your idea really well, penetrate social networks that govern the usage of a medication strategy and employ long-term education efforts.

Even then it can take many decades for a new idea to get used.

Human cognitive biases, ego and financial incentives appear to have played a role in halting past adoption of medical innovations.

Humanity's track record of adopting innovation has not been exemplary with the cost of millions of lives as the puerperal fever and scurvy examples make clear.

Humans Very Reluctant to Cede Control to Machines

In general, humans have been very reluctant to cede control of their activities to machines. Think about the nervousness associated with self-driving cars as an example.

But in the case of medicine one wonders if the greater use of AI in diagnosis of medical conditions and the delivery of medicine itself might accelerate the adoption of innovative ideas while improving the practice of medicine overall.

There has been a lot written in this area and there are quite a few good reasons to be optimistic about the potential for AI in the practice of medicine.

Interestingly, we saw a recent talk from the former head of the AMA who was quite negative on allowing AI into medicine but our own feeling was that the speaker was not fully seeing the potential benefits.

The speaker blamed 737 plane crashes on AI and wondered aloud if we wanted to allow AI anywhere near a medical office. It felt almost like listening to New York cab drivers argue that Uber would be bad for customers and that Uber drivers might drive customers into the river etc.

Will AI Make A Difference in Adoption of Innovation?

The Access and Rationality Argument

Today's LLM's combine access to massive amounts of scientific data and reasoning engines to attempt to the find the right answer to an external query.

Overall, the ability of these systems to find the right answer is easily beating what doctors can do and medical consumers are adopting chatbots at a rapid pace.

Suppose that Semmelweis published a paper on puerperal fever on medRxiv last month with a convincing set of empirical supportive data. Further, suppose that existing miasmatic theories had no data but did have strong support from existing opinion leaders in obstetrics.

Because reasoning engines enforce rationality over opinion it would be quite likely that medical AI systems would recommend Semmelweis' approach to preventing puerperal fever even though the explanation for these results might not have been fully elucidated.

Efficacy Signals Science Access and Diffusion

The fact that numerous studies show that LLM / reasoning systems outperform physicians in diagnostic tasks indicate that these computerized systems are doing a better job of accessing the right information at the right time than physicians do.

This can be seen as prima facie evidence that LLM-based systems are highly likely to rely on evidence-based recommendations.

The increasing use of real-time document access (called retrieval-augmented generation) in LLM's with rigorous reasoning engines further highlights the relatively unbiased way in which LLM's access the right science to make medical recommendations.

Interestingly, physicians are increasingly using LLM's themselves in clinical practice which means that these new systems are able to help with the diffusion of new clinical knowledge.

Will AI Make A Difference in Adoption of Innovation?

Online Publishing / Open Science Model

Obviously, online publishing is itself quite important because the body of knowledge of what matters in medicine can be captured by computers and analyzed using LLM type systems without human curation.

Open science is a movement to make research—including publications, data, samples, and software—freely accessible, transparent, and collaboratively developed across all levels of society.

Importantly, the open science model that today has allowed scientists to release preprints of their papers online without peer review reduces the opportunity for the existing "medical establishment" – whatever that might be, to suppress new ideas.

This is not an unalloyed good thing. Obviously, the absence of peer review makes it easier to publish false results, pseudoscience and the like.

Pressure of the Healthcare Consumer

Another key factor with AI is that it is allowing many healthcare consumers to disintermediate their physicians.

Consumers can often figure out their diagnoses on their own using LLM systems and may be less prone to consult physicians.

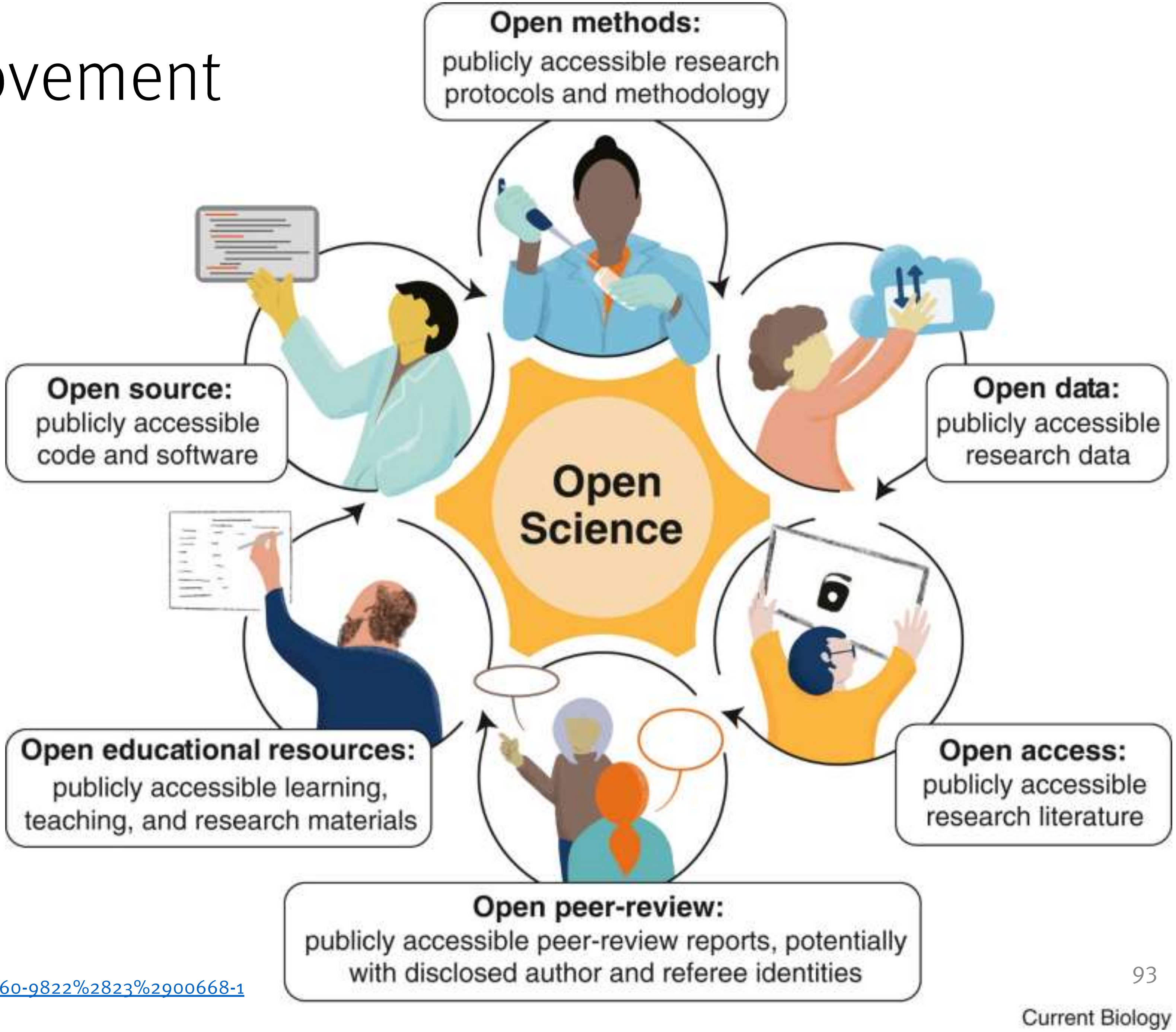
Further, consumers can check the advice their doctor is giving them against an LLM.

This puts obvious pressure on doctors to give the best advice knowing that a competitive source of medical information and advice is available.

Doctors care about this because they make money from their customers and so want to make them as happy as possible.

The Open Science Movement

The open science movement is designed to make scientific research (including publications, data, physical samples, and software) and its dissemination accessible to all levels of society, amateur or professional. Open science is transparent and accessible knowledge that is shared and developed through collaborative networks. It encompasses practices such as publishing open research, campaigning for open access, encouraging scientists to practice open-notebook science (such as openly sharing data and code), broader dissemination and public engagement in science and generally making it easier to publish, access and communicate scientific knowledge.



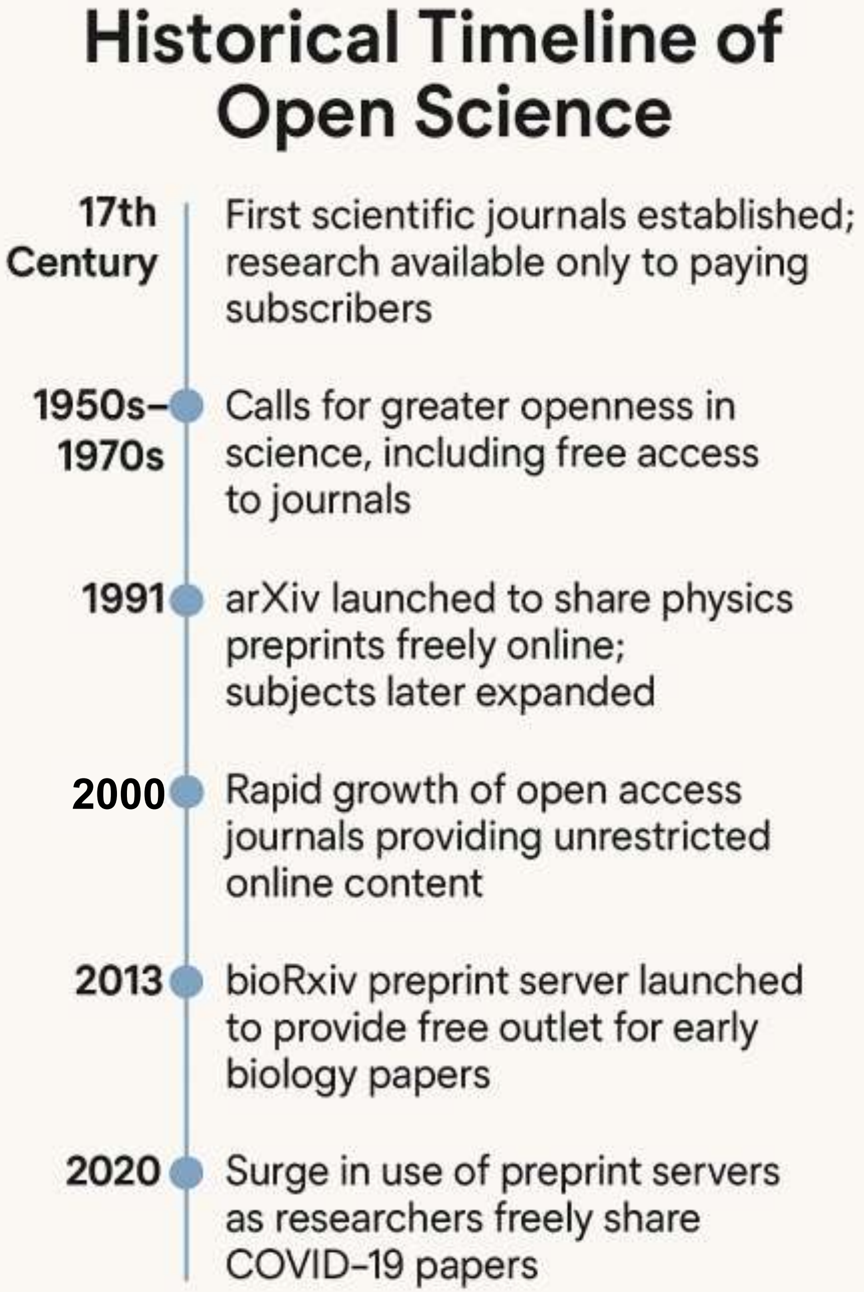
Source of image: https://www.cell.com/current-biology/fulltext/S0960-9822%2823%2900668-1

Open Science Journals Accelerate Diffusion of Knowledge

Misha Teplitskiy, Grace Lu and Eamon Duede, *arXiv*, June 25, 2015

With the rise of Wikipedia as a first-stop source for scientific knowledge, it is important to compare its representation of that knowledge to that of the academic literature. Here we identify the 250 most heavily used journals in each of 26 research fields (4,721 journals, 19.4M articles in total) indexed by the Scopus database, and test whether topic, academic status, and accessibility make articles from these journals more or less likely to be referenced on Wikipedia. We find that a journal's academic status (impact factor) and accessibility (open access policy) both strongly increase the probability of it being referenced on Wikipedia. Controlling for field and impact factor, the odds that an open access journal is referenced on the English Wikipedia are 47% higher compared to paywall journals. One of the implications of this study is that **a major consequence of open access policies is to significantly amplify the diffusion of science**, through an intermediary like Wikipedia, to a broad audience.

Source: https://arxiv.org/abs/1506.07608



Open Science Made a Difference in Dealing with the Covid-19 Epidemic

Fraser N, Brierley L, Dey G, Polka JK, Pálfy M, Nanni F, Coates JA. The evolving role of preprints in the dissemination of COVID-19 research and their impact on the science communication landscape. *PLoS Biol.* 2021 Apr 2;19(4):e3000959.

The world continues to face a life-threatening viral pandemic. The virus underlying the Coronavirus Disease 2019 (COVID-19), Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2), has caused over 98 million confirmed cases and 2.2 million deaths since January 2020. Although the most recent respiratory viral pandemic swept the globe only a decade ago, the way science operates and responds to current events has experienced a cultural shift in the interim. The scientific community has responded rapidly to the COVID-19 pandemic, releasing over 125,000 COVID-19–related scientific articles within 10 months of the first confirmed case, of which more than 30,000 were hosted by preprint servers. We focused our analysis on bioRxiv and medRxiv, 2 growing preprint servers for biomedical research, investigating the attributes of COVID-19 preprints, their access and usage rates, as well as characteristics of their propagation on online platforms. Our data provide evidence for increased scientific and public engagement with preprints related to COVID-19 (COVID-19 preprints are accessed more, cited more, and shared more on various online platforms than non-COVID-19 preprints), as well as changes in the use of preprints by journalists and policymakers. We also find evidence for changes in preprinting and publishing behaviour: COVID-19 preprints are shorter and reviewed faster. Our results highlight the unprecedented role of preprints and preprint servers in the dissemination of COVID-19 science and the impact of the pandemic on the scientific communication landscape.

Source: https://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.3000959

Medicine To Become Even More Evidence-Based and Adopt Innovations More Quickly

We began this report by noting historical foibles in knowledge transmission and innovation adoption.

There has been clear progress with introduction of the Web and, particularly now, with AI tools like ChatGPT.

Current developments on vaccines notwithstanding, in the long run we have seen greater reliance on scientific evidence in implementation of medical interventions. Human lifespans are getting longer, showing the positive effects of numerous positive developments in public health and medicine. Reductions in maternal and infant mortality in the last century have been particularly impressive.

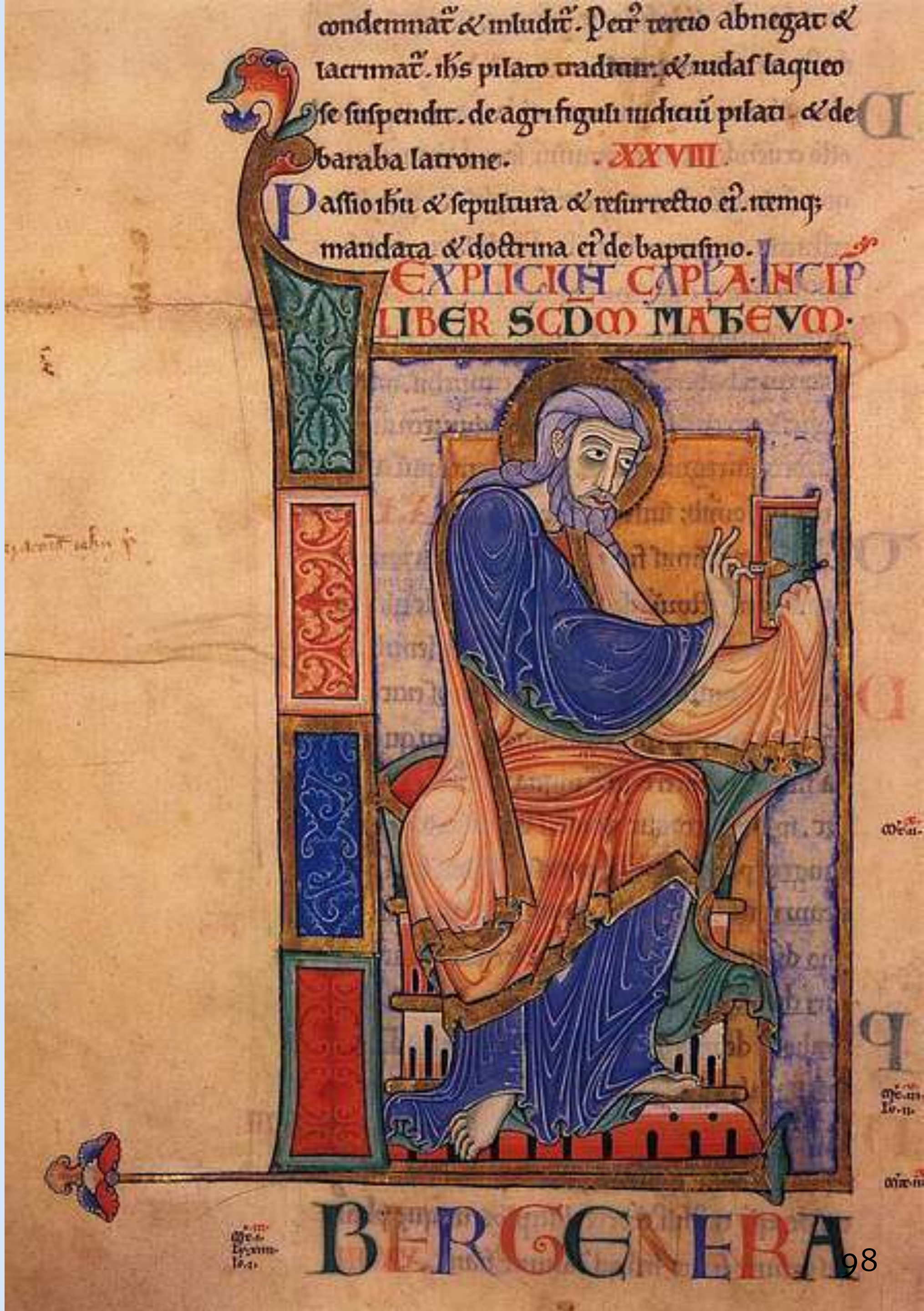
Two very important recent trends promise to accelerate medical progress and make it less likely that the next great medical discovery gets to patients in a timely way. One trend is the open science movement – which is accelerating the time from discovery to release by allowing scientists to bypass the time-consuming review and editorial processes associated with traditional journals. The second trend is the LLM combined with the reasoning engine. This has led to meaningful change in primary care medicine in just a few years and promises far more change in the future. Ultimately, AI tools should be able to do a far better job of handling the onslaught of proteomics and other data to design improved medical interventions – with positive implications for our species for centuries to come.

Appendix: Accessing Stifel Publications and Events

Past Issues

To get on the mailing list for this publication feel free to contact Jenna Hill (hillje@stifel.com). Past issues of this publication can be read online at:

Aug 18, 2025 (Cardiovascular Drugs)
Jul 14, 2025 (Top 40 Pharma)
Jun 23, 2025 (Science and Truth)
May 12, 2025 (MFN Policy)
May 5, 2025 (NIH Cuts, China Tariffs)
Apr 28, 2025 (Eyes on Washington DC)
Apr 21, 2025 (FDA Shifts, Buyside Update)
Apr 14, 2025 (Wild Week in Market)
Apr 7, 2025 (Biotech Market Break)
Mar 31, 2025 (China Biotech Update)
Mar 24, 2025 (Healthcare Reform)
Feb 24, 2025 (Retail Pharma Trends)
Feb 10, 2025 (Pharma Earnings)
Jan 27, 2025 (Women's Health, Obesity)
Dec 17, 2024 (Biotech Blues)
Nov 25, 2024 (Biotech Balance Sheets)
Nov 18, 2024 (New Administration)
Nov 4, 2024 (Election, Obesity)
Oct 21, 2024 (China, Pfizer)
Oct 7, 2024 (VC update)
Sep 23, 2024 (The Fed Rate Cut)
Sep 9, 2024 (Sector Outlook)
Aug 12, 2024 (Biotech Market)
July 8, 2024 (Obesity Market Update)
June 17, 2024 (Lab Market)
June 8, 2024 (Oncology Review)
May 27, 2024 (GLP-1's)
May 20, 2024 (Returning Capital)
May 13, 2024 (Brain, AlphaFold 3)
May 6, 2024 (Earnings, Obesity)
April 29, 2024 (M&A, Japan)
April 22, 2024 (Pharma Pricing)
April 15, 2024 (AI in Pharma)
April 8, 2024 (The Buyside)
April 1, 2024 (Biotech Balance Sheets)
March 25, 2024 (Women's Health)
March 18, 2024 (Inflammasome)
March 11, 2024 (IRA, Immunology)
March 4, 2024 (Biotech Employment)
Feb 26, 2024 (Biotech Strategy)
Feb 19, 2024 (Big Drugs, Autoantibodies)
Feb 12, 2024 (Fibrosis, Endometriosis)
Feb 5, 2024 (Severe Disease in Women)
Jan 29, 2024 (Pharma R&D Productivity)
Dec 18, 2023 (Expectations for Future)
Dec 11, 2023 (ASH, R&D Days)
Dec 4, 2023 (Big Pharma, CEA)
November 20, 2023 (M&A)
November 13, 2023 (AHA, Bear Market)
November 7, 2023 (Unmet Needs)
October 30, 2023 (ADCs)
October 23, 2023 (ESMO Review)
October 16, 2023 (Cancer Screening)
October 9, 2023 (Biosimilars, M&A)
October 2, 2023 (FcRn, Antibiotics)
September 25, 2023 (Target ID)
September 18, 2023 (Pharma Strategy)
September 11, 2023 (US Health System)
September 5, 2023 (FTC, IRA, Depression)



12th-century painters - Dover Bibl, Parker Library, Corpus Christi College, Cambridge.

Links to Stifel Biopharma Special Topic Publications

Obesity Drug Update

July 9, 2025

Oncology Update

Jun 5, 2025

Aging Biology, Part I

Mar 26, 2025

2025 Biotech Outlook

Jan 8, 2025

2024 Biotech Mid-Year Outlook

July 15, 2024

Obesity Drug Update

July 8, 2024

AI in medicine

Jan 22, 2024

Why Invest in Biotech?

November 22, 2023

Feel Free to Join Us at Biotech Hangout

Please join us this Friday at noon EST for the latest episode.

To Learn More
https://www.biotechhangout.com/

Disclosure

Stifel collectively refers to Stifel, Nicolaus & Company, Incorporated and other affiliated broker-dealer subsidiaries of Stifel Financial Corp. The information and statistical data contained herein have been obtained from sources that Stifel believes are reliable, but Stifel makes no representation or warranty as to the accuracy or completeness of any such information or data and expressly disclaims any and all liability relating to or resulting from your use of these materials. The information and data contained herein are current only as of the date(s) indicated, and Stifel has no intention, obligation, or duty to update these materials after such date(s). These materials do not constitute an offer to sell or the solicitation of an offer to buy any securities, and Stifel is not soliciting any action based on this material. Stifel may be a market-maker in certain of these securities, and Stifel may have provided investment banking services to certain of the companies listed herein. The author of this report, Tim Opler, holds ownership stakes in a number of private and public healthcare companies including eGlint and Health Outlook, both mentioned in this report. Stifel and/or its respective officers, directors, employees, and affiliates may at any time hold a long or short position in any of these securities and may from time-to-time purchase or sell such securities. This material was prepared by Stifel Investment Banking and is not the product of the Stifel Research Department. It is not a research report and should not be construed as such. This material may not be distributed without Stifel's prior written consent.

Stifel, Nicolaus & Company, Incorporated | Member SIPC & NYSE | www.stifel.com